Applied Physics B (2021) 127:5
https://doi.org/10.1007/s00340-020-07564-8

Applied Physics B
Lasers and Optics

Research on the measurement of CO_2 concentration based on multi-band fusion model

Honglian Li[1] · Shuai Di[1,2] · Wenjing Lv[1] · Yaqing Jia[1] · Shijie Fu[3] · Lide Fang[1]

Received: 13 June 2020 / Accepted: 8 December 2020 / Published online: 2 January 2021

Abstract
In this paper, a multi-band fusion model to improve the performances of the super-continuum laser absorption spectrometer (SCLAS) of CO_2 was proposed and demonstrated. Various concentrations of CO_2 were measured by the super-continuum laser in the wavelength of 1425–1445 nm, 1565–1585 nm, and 1595–1615 nm, respectively, at 295 K and 1 atm. The method for derivation of CO_2 concentration using the integrated area of spectrum peaks is proposed. Linear models of the CO_2 concentration and the integrated area of the absorption peaks in different bands were established, which achieves R^2 of 0.9947, 0.9937, and 0.9824, respectively. The measurement accuracy with the models is evaluated with the parameter of relative analysis error (*RPD*), which results in a of more than 2, indicating reasonable prediction ability of the models. In order to improve the accuracy of the single model, the models of the three bands are weighted and fused based on R^2 and *RMSE*, respectively. One of the fusion models reduces the prediction error and improves the accuracy of the single model effectively by decreasing the maximum relative error from 3.4 to 1.2% for a single model. The experimental results show that SCLAS can measure the CO_2 concentration under different environments. The multi-band fusion model proposed here is feasible for CO_2 measurement, which provides a new idea and new method for the detection of gases.

1 Introduction

CO_2 is the primary greenhouse gas in the atmosphere, and about 60% of the greenhouse effect is caused by CO_2. With the rapid development of the industry, the problem of air pollution has become an increasingly prominent issue [1, 2]. The increase of carbon emissions not only exacerbates the greenhouse effect but also has a serious impact on the daily life of human beings. Therefore, the accurate measurement of atmospheric CO_2 concentration is of great importance in improving the scientific understanding of the CO_2 impacts, which could also provide guidance on the control of carbon emission.

The methods used to determine CO_2 concentrations currently can be divided into chemical methods and spectrum measurement methods. The methods of spectral measurement have been widely used in environmental detection because of its wide detection range, high sensitivity, and real-time online analysis. The absorption spectrum detection technique is an effective detection technique with high sensitivity, wide detection range, and strong practicability, which has a good application prospect in the field of gas detection. Infrared spectroscopy technologies commonly used include Tunable diode laser absorption spectroscopy (TDLAS) [3], Fourier transform infrared spectroscopy (FTIS) [4], Photoacoustic spectroscopy (PAS) [5]. Compared with these three technologies, super-continuum laser detection technology is a new type of absorption spectrum detection technology, which can not only measure the single gas component but also measure the multi-component gases in the atmosphere simultaneously [6]. Super-continuum laser has become an ideal light source for detecting gases for its easy collimation, high brightness, wide spectrum and high stability [7–9]. Jihyung et al. [10] have measured low concentrations of hydrocarbon gases such as acetylene and ethane by super-continuum laser detection technology. The experimental results were consistent with the results of HITRAN database. It is verified that the super-continuum laser can detect broadband absorption spectra. Wen et al. [11] proposed a multi-model

✉ Honglian Li
lihonglian@hbu.edu.cn

[1] College of Quality and Technical Supervision, Hebei University, Baoding Hebei 071002, China

[2] College of Chemistry and Environmental Science, Hebei University, Baoding Hebei 071002, China

[3] College of Optical Sciences, University of Arizona, 1630 E. University Blvd, Tucson, AZ 85721, USA

fusion variable selection method for the characteristics of near-infrared spectroscopy data, which improved the prediction ability of the model and verified the feasibility of the method experimentally. Pan et al. [12] used the spectroscopy method to detect the total nitrogen concentration in water, in which the data of a single model of different bands was fused. The established fusion model improved the system measurement accuracy effectively. To improve the reliability and stability of the mine environmental monitoring system, Wang et al. [13] used the weighted data fusion technology to process the data of multiple sensors to decrease the measurement error of gas concentration. Tobias et al. [14] measured N_2O and CO_2 in cultivated land, forests, and grassland through a new data fusion model, which helped to control the gas emissions. In order to perform the high-level fusion for quantitative analysis, Li et al.[15] proposed two MID methods of *RMSEPW* and *RPDW* to proceed with high-level fusion.

In summary, the application of fusion models improves the measurement accuracy of the detection systems, but researches using the super-continuum laser absorption spectroscopy technology combined with the fusion method on gas concentration detection have rarely been reported. To reduce the system measurement error and improve the accuracy of CO_2 prediction, this paper proposes a multi-band weighted fusion model for CO_2 concentration measurement. The prediction effect is better than other single-band prediction models and it overcomes the shortcomings of single model prediction, also it is easy to implement and it has strong applicability for the CO_2 measurement. This study takes CO_2 as an example and set up a super-continuum laser absorption spectrum detection system. The absorption spectrum of CO_2 are measured in different bands. The method of inversion measurement of gas concentration through the integrated area of absorption peaks is proposed, and the measurement model of CO_2 concentration in different bands is established. The fusion model obtained based on *RMSE* improves the accuracy of the single model effectively by decreasing the maximum relative error from 3.4 to 1.2% for a single model. This study verifies that the multi-band fusion model is feasible in the measurement of gas concentration, which provides a new method and new idea for the concentration measurement of other gases in the future.

2 Theoretical basis

2.1 Multi-band fusion model

To improve the predictive ability of a single model, weights are assigned to the three measurement models to obtain a multi-band fusion model, thereby improving the system measurement accuracy, which is of great significance for actual environmental monitoring. The fusion models obtained by weighted fusion of the multiple single models are Z_G, and it can reduce the interference of water vapor and other gases in the air on the experimental results effectively. The construction of the fusion model is shown in Fig. 1. The model can be described as:

$$Z_G = \sum_{i=1}^{n} Y_i W_i = (i = 1, 2...n). \tag{1}$$

In Eq. (1), ZG (G = A, B…G) is the fusion model, G is the number of the fusion model, n is the number of the single model, Yi is the single model. The weight W is introduced into the model, $Wi \in [0,1]$, (i = 1,2…n). The weight represents the relative importance of a single model in the fusion model. The most significant step in establishing the fusion model is to determine the weight coefficient of each single model.

In this paper, the weight distribution of the three models is evaluated by the fitting coefficient R^2 and the mean square error *RMSE* of each single model. The fusion model is used to reduce the prediction error and improve the accuracy of the concentration inversion.

2.2 Weighting indicator determined method

2.2.1 Fusion model Z_G based on R^2

R^2 refers to how well the linear model fits the values observed. The closer the value of R^2 to 1, the better the model fits. The weight coefficients of the fusion model Z_G are determined by the fitting coefficients R^2 of the three models. The weight is determined by the proportion of R^2 of a single model in the sum of R^2 of three models. The sum of the fitting coefficients e_1 of the single model can be

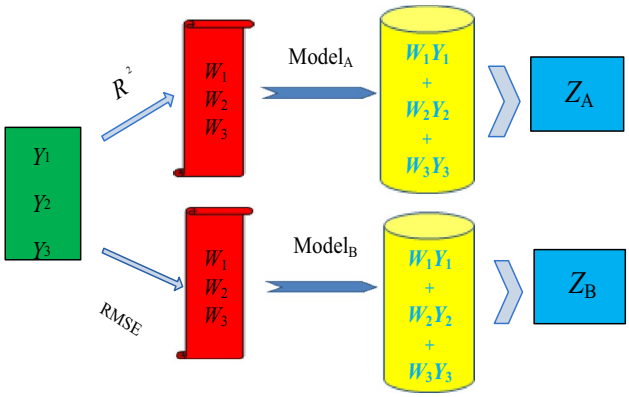

Fig. 1 The construction of the multi-band weighted fusion model. The new fusion models Z_A and Z_B are obtained and the weights are determined by R^2 and *RMSE*, respectively

obtained through experimental results. The e_1 and W_i are shown below:

$$e_1 = \sum_{i=1}^{n} R_i^2 (i = 1, 2...n), \tag{2}$$

$$W_i = R_i^2 / e_1 \quad (i = 1, 2...n), \tag{3}$$

2.2.2 Fusion model Z_G based on *RMSE*

Root mean square errors (*RMSE*) is used to measure the deviation between the observed value and the true value. It is often used to evaluate the prediction result of the model. The weight coefficients of the fusion model Z_B are derived from the mean square error *RMSE* of the three single models [16]. The e_2 is the sum of the *RMSE* of the single model, also it can be derived from the experimental data, M is the number of single models. The e_2 and W_i can be calculated as:

$$e_2 = \sum_{i=1}^{n} \mathrm{MSEi} \quad (i = 1, 2...n), \tag{4}$$

$$W_i = \frac{e_2 - \mathrm{MSE}_i}{e_2} * \frac{1}{M-1} \quad i = 1, 2...n, \tag{5}$$

2.3 Selection and analysis of absorption lines

The super-continuum laser selected has a spectral range of 400–2400 nm, and the wavelength range filtered by a Laser Line Tunable Filter (LLTF) is 1000–1700 nm. To improve the sensitivity and signal-to-noise ratio of the spectrometer and to enhance the absorption effect of CO_2, a wavelength band with strong absorption was selected in the experiment. We should evaluate the spectral interference generated by the absorption line of atmospheric gases to ensure the measurement precision. To further avoid the peak and trough of the background spectrum, the wavelength of 1280–1700 nm is selected for the detection. According to the HITRAN2016 gas spectrum database, the absorption spectrum of CO_2 within the range 1280–1700 nm (5882–7813 cm^{-1}) is shown in Fig. 2.

It can be seen that there are six CO_2 absorption clusters lying between 1425 and 1615 nm. Therefore, the multi-band measurements of CO_2 were performed in three wavelength bands which are 1425–1445 nm, 1565–1585 nm, and 1595–1615 nm, respectively. It is well known that the characteristic absorption spectra of gas molecules are decided by their own structure. To avoid the interference of other gases on the CO_2 absorption line, the absorption line distributions of H_2O and the other gases in the air are compared with CO_2

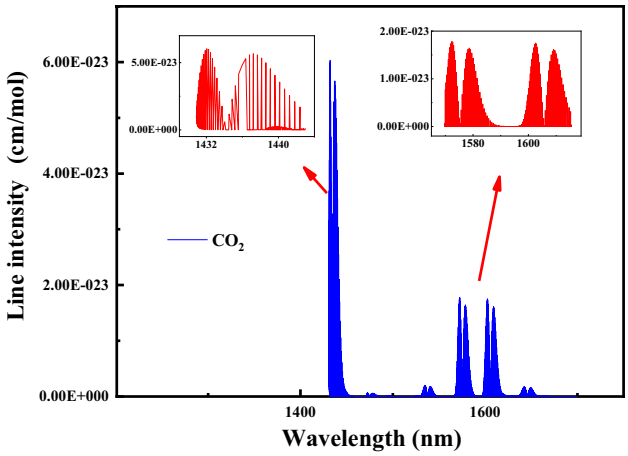

Fig. 2 Absorption spectrum of CO_2 within the range 1280–1700 nm at 1 atm, 295 K. Data obtained from HITRAN2016 database

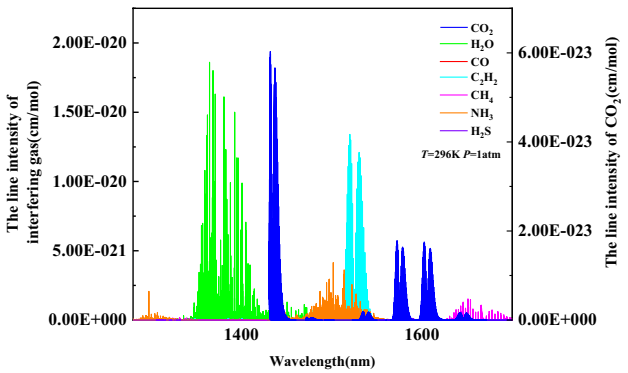

Fig. 3 The comparison of absorption spectra of H_2O, CO, H_2S, C_2H_2, NH_3 and CH_4 with CO_2 in 1280–1700 nm at 1 atm and 296 K based on the HITRAN2016 database

at 1280–1700 nm at 1 atm based on the HITRAN2016 database. The absorption intensity of HF, O_2 and O_3 in this band is relatively weak, and their influence on CO_2 measurement is negligible. SF_6, SO_2, C_2H_4, C_2H_6, HCHO, CH_3OH, NO_2 and other gases have no absorption spectrum in this band. The possible interference from H_2O, CO, H_2S, C_2H_2, NH_3, CH_4, and other gases on the CO_2 absorption line in this band is shown in Fig. 3. The absorbance of CO_2 and H_2O, CO, C_2H_2, CH_4, H_2S, NH_3 at different concentrations obtained from the database at the bands of 1425–1445 nm, 1565–1585 nm, and 1595–1615 nm are shown in Fig. 4.

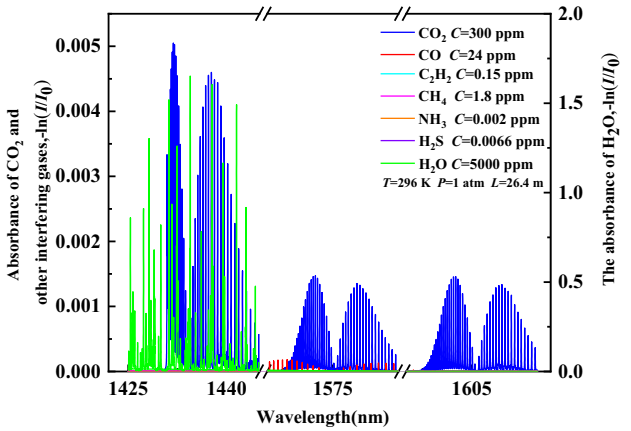

Fig. 4 The comparison of absorbance of H_2O, CO, C_2H_2, CH_4, H_2S and NH_3 with CO_2 in 1425–1445 nm, 1565–1585 nm and 1595–1615 nm at 1 atm, 296 K and 26.4 m optical path based on the HITRAN2016 database

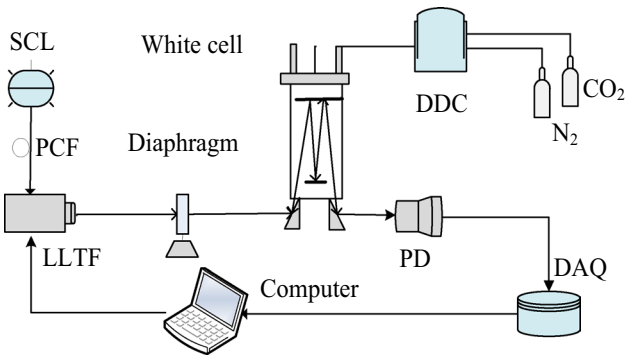

Fig. 5 Schematic of the experimental setup. *SCL* Super-continuum laser, *PCF* Photonic crystal fiber, *LLTF* Laser Line Tunable Filter, *DDC* Dynamic dilution calibrator, *PD* Photo-detector, *DAQ* Data acquisition card

3 Experiments

3.1 Experimental system

Our experimental SCLAS consists of a super-continuum laser, a Laser Line Tunable Filter (LLTF), a diaphragm, a dynamic dilution calibrator, a White cell, a photo-detector, a data acquisition card (DAQ), and a personal computer (PC). The experimental setup is shown in Fig. 5. The super-continuum laser (type SC400-4, Fianium, UK), a picosecond pulsed and wavelength between 400 and 2400 nm with a maximum output power of 8 W, was used and wavelength filtered by the LLTF, and programing was controlled by the PC, so that the output laser wavelength is scanned in the wavelength range of 1425–1445 nm, 1565–1585 nm, 1595–1615 nm. After filtering out the stray light around the beam, the laser incidents in the White cell. It is then reflected in the White cell multiple times and reaches the photo-detector. The photo-detector converts the optical signal into an electrical signal and transmits the data to the DAQ and PC terminals.

The LLTF is used in the experiment, which is the light source's spectroscopic accessory in this system. The spectral bandwidth is 5 nm, and the wavelength tuning resolution is 0.1 nm. The output spectral range is 1000–2300 nm, which has a high damage threshold and long lifetime.

The White cell (model 35-V-H, Infrared Analysis, USA), with a multipass long path of 2.2–35 m and volume 8.5 L, was used to improve the sensitivity of the spectrometer. The incident laser is reflected multiple times in the cell to increase the optical path length, thereby achieving the purpose of gas circulation absorption. The dynamic dilution calibrator employs two mass flow meters with high-precision to control the gas flow rates, thus obtaining the desired concentration of the test gas. The flow measurement accuracy is ± 1.0%, the flow control repeatability is ± 0.2%, and the flow measurement linearity is ± 0.5%.

The diaphragm is a GCM-5711 M variable square aperture diaphragm, which can filter out the surrounding stray light, thereby adjusting the intensity of the passing beam. The photo-detector (PDA50B, THORLABS, USA), with a spectral range of 800–1800 nm, a gain of 0–70 dB and response time of 50 ns, is used to detect the characteristic light signal absorbed by the gas with low noise.

The experiment was carried out at room temperature (295 K) and at the gas pressure of 1 atm. The light path was aligned by visible light before the experiment to ensure that the laser could be smoothly injected from the left side of the white cell and emitted from the right side. The laser was reflected back and forth in the cell for 21 times and the corresponding optical path reached 26.4 m. During the experiment, high-purity nitrogen (N_2) (99.999%) was first introduced into the absorption cell and the measured signal was taken as the background signal. Afterwards, 5.0%, 5.5%, 6.5%, 7.0%, 7.5% and 8.0% of CO_2 were injected in the absorption cell for testing. In order to ensure the accuracy of the experimental results, the absorption cell was purged with N_2 for ten minutes at 1 atm before each measurement.

4 Results and analysis

4.1 Measurement results

The absorption spectrum of CO_2 at different concentrations tested at the bands of 1425–1445 nm, 1565–1585 nm, and 1595–1615 nm are shown in Fig. 6(a)–(c). The background subtraction method is used to eliminate the influence of low-frequency background noise such as the

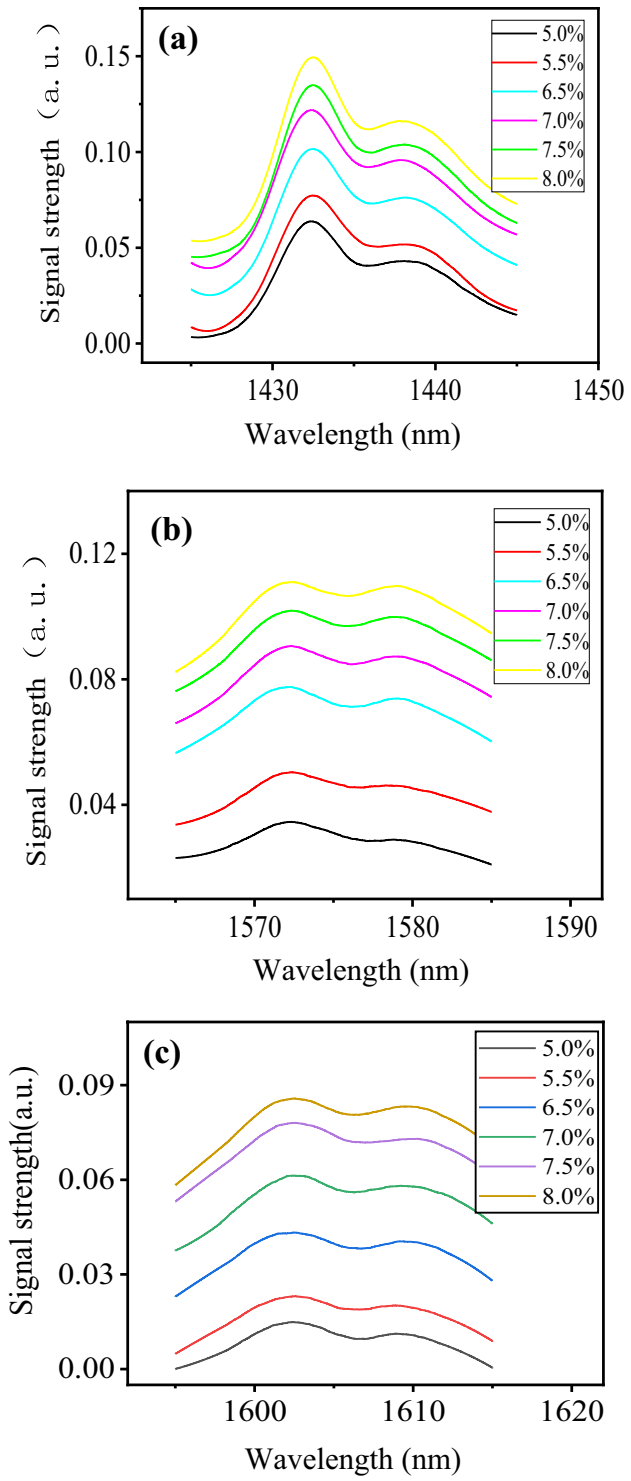

Fig. 6 The absorption spectrum of CO_2 with different concentrations at (**a**) 1425–1445 nm, (**b**) 1565–1585 nm, and (**c**) 1595–1615 nm at 295 K and 1 atm

baseline drift. The data is smoothed by the S–G filtering method, and then the sampling points and wavelengths are converted from a time domain to a frequency domain. The experimental results show that in all the three tested bands, there appears to be main and secondary absorption peaks, which are 1432 nm/1437 nm, 1572 nm/1579 nm, and 1603 nm/1609 nm, respectively. The results are consistent with those in the database. The absorbed signal intensity has a good positive correlation with gas concentration. As the gas concentration increases, the gas absorption signal gradually increases.

4.2 Modeling and evaluation

However, the intensity of the gas absorption peak may be deviant from the real value due to the drift of the line center, the spectral line type, and the spectral line broadening. Therefore, this paper proposes a method for the derivation of gas concentration through the integrated area of the CO_2 absorption peak. The method can eliminate the influence of single-wavelength light intensity fluctuation and spectral line wavelength shifting effectively, and avoid interference of other factors such as noise. Three linear models Y_1, Y_2 and Y_3 are established for the absorption bands 1430–1435 nm, 1570–1575 nm, and 1600–1605 nm respectively, which reveals the relationship between the integral area and the CO_2 concentration of the main absorption peak. The models are depicted in Fig. 7(a)–(c).

The prediction accuracy of the three models is evaluated by the determination coefficient R^2 of the actual value and the predicted value, the root mean square error (*RMSE*), and the relative analysis error (*RPD*). The larger the R^2, the smaller the *RMSE*, indicating the higher the accuracy of the model. *RPD* [17, 18] is defined as $RPD = SD/RMSE$, where *SD* is the sample standard deviation. When $RPD \geq 2.0$, it indicates that the model is suitable for prediction. When $1.4 < RPD < 2.0$, the model is considered to be more reliable, and the prediction accuracy of the model can be improved by other modelling methods. When $RPD \leq 1.4$, the model is considered to be unreliable. The estimation accuracy of the three models is shown in Table 1.

The experimental results show that there is a good positive correlation between the integral area of the absorption peak and the gas concentration. The larger the CO_2 concentration, the larger the integrated area of the peak. The *RPD* of the three concentration measurement models is all greater than 2, indicating that the model has high stability and strong prediction ability in deriving the CO_2 concentration. However, the strong absorption lines of other gases in the air will interfere with the CO_2 absorption line thus affecting the predicting results of the models. Therefore, to reach the best prediction results, the three models have to be used according to different situations.

The model Y_1 has a higher fitting coefficient and the highest stability among the models. However, the measurement band of the model is greatly affected by H_2O, NH_3 and

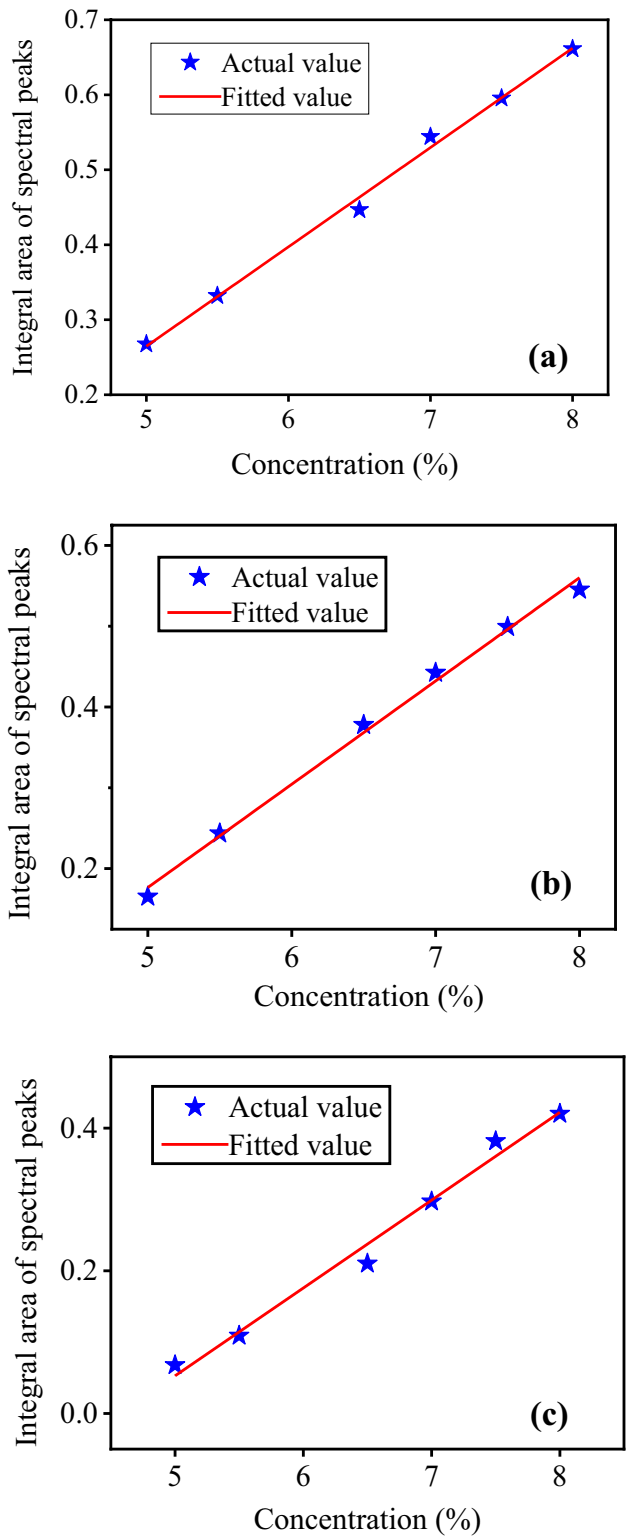

Fig. 7 The linear fitting results of the integrated area and concentration at the main absorption peaks of (**a**) 1430–1435 nm, (**b**) 1570–1575 nm, and (**c**) 1600–1605 nm at 295 K and 1 atm

CH_4, thus the concentration of CO_2 can be measured in the absence of these three gases. When Model Y_2 is employed, it is interfered with by the absorption lines of CO, C_2H_2, NH_3, H_2S, etc., thus the CO_2 concentration can be measured in the absence of these gases. Model Y_3 is only greatly affected by the CH_4 and H_2S. Compared with the first two models, the model Y_3 is less interfered by other gas types. When the effect of CH_4 and H_2S in the air is small, the model Y_3 can be used to detect the CO_2.

The weights were calculated based on the R^2 and *RMSE* of the single model Y_1, Y_2, Y_3, and the results are shown in Fig. 8.

The maximum relative errors of the two fusion models obtained through data processing are shown in Table 2. After analysis, it can be found that the least-squares method with weight can improve the model inversion ability effectively, and the multi-band fusion model can reduce the measurement error greatly, and thus achieving the purpose of improving the concentration measurement accuracy. Comparing with the single model, the fusion models obtained by R^2 and *RMSE* have both reduced the measurement error, and the fusion model Z_B reduces the error greatly, which improves the accuracy of the measurement results effectively. In summary, the fusion model Z_B overcomes the problem of the low prediction accuracy of a single model, therefore, this fusion model can be selected to measure the CO_2 concentration.

Comparing with the CO_2 concentration measurement results reported currently, the maximum concentration measurement error is 1.2%, which is improved the prediction accuracy effectively. Yao et al. [19] used DAS and wavelength modulation spectroscopy (WMS) to measure the CO_2 concentration, while the maximum relative errors for DAS and WMS were 2.64% and 1.65%, respectively. Deng [20] reported that they used a distributed feedback (DFB) diode laser as the local oscillator to measure atmospheric CO_2 column concentration continuously in the near-infrared region, the averaged measurement precision is 1.6% by analyzing the standard deviation of CO_2 column concentration. Yang [21] used the Li-7500 analyzer to calibrate and analyze the miniaturized atmospheric CO_2 detection system. the absolute value of the relative error of carbon dioxide volume ratio was less than 2.0% by the inversion.

5 Conclusion

In this study, a SCLAS with a multi-band fusion algorithm was proposed and demonstrated. The absorption spectra of CO_2 in different bands of near-infrared were measured. Three linear models of the integrated absorption peak area and the CO_2 concentration in different bands are established, and the stability of the model is evaluated. Two new fusion

Table 1 Accuracy assessment of model Y_1, Y_2 and Y_3

Waveband (nm)	Model	R^2	*RMSE*	*SD*	*RPD*
1430–1435	$Y_1=0.1325X_1-0.3977$	0.9947	0.0112	0.1538	13.73
1570–1575	$Y_2=0.1278X_2-0.4624$	0.9937	0.0118	0.1484	12.58
1600–1605	$Y_3=0.1232X_3-0.5631$	0.9824	0.0191	0.1437	7.52

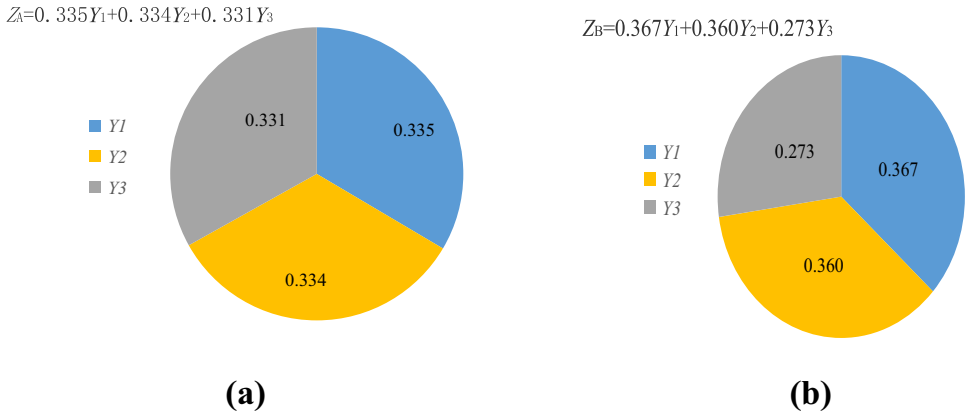

Fig. 8 The pie chart of the weight of a single model. (**a**) Weighting indicator based on R^2; (**b**) weighting indicator based on *RMSE*

Table 2 Relative error analysis of concentration inversion for each model

Model *RE*	Single-band model (%)	Fusion model Z_A (%)	Fusion model Z_B (%)
Band (nm)			
1430–1435	2.0	2.0	1.2
1570–1575	1.8		
1600–1605	3.4		

models are obtained by assigning weights to the three linear models based on R^2 and *RMSE*, respectively. It is found that the fusion model obtained based on *RMSE* is better, and the relative error of CO_2 concentration inversion is reduced to 1.2%, which improves the accuracy of the model effectively. The experimental results show that it is feasible to measure the CO_2 concentration by a multi-band fusion model, which provides a new idea and new method for the detection of other gases and multi-component gases.

Acknowledgements This work has been supported by the Key Projects of Hebei Natural Science Foundation (No. E2017201142); Hebei Natural Science Youth Fund (No. D2012201115); 2018 Ministry of Education "Chunhui Program" Cooperative Scientific Research Projects; the Postdoctoral Research Projects in Hebei Province (No. B2016003008).

References

1. J.Y. Chen, Infrared Phys. Technol. **80**, 131–137 (2017)
2. T. Cai, G. Gao, M. Wang et al., J. Quant. Spectrosc. Radiat. Transfer **201**, 136–147 (2017)
3. Y.D. Wang, Sens. Actuat. B **225**, 188–198 (2016)
4. L. Liu, A. Mandelis, A. Melnikov et al., Int. J. Thermophys. **37**(7), 64 (2016)
5. Q. Wang, Z. Wang, J. Chang et al., Opt. Lett. **42**(11), 2114 (2017)
6. E. Johannes, Sci. Rep. **8**, 10312 (2018)
7. T. Werblinski, S.R. Engel, R. Engelbrecht et al., Opt. Express **21**(11), 13656 (2013)
8. P.S. Edwards, M.D. Turner, G.W. Kamerman et al., SPIE Proc. **7323**, 73230S (2009)
9. C. Amiot, A. Aalto, P. Ryczkowski et al., Appl. Phys. Lett. **111**(6), 061103 (2017)
10. Y. Jihyung, Appl. Spectrosc. **70**(6), 1063–1071 (2016)
11. M.J. Hong, Z.Y. Wen, Spectrosc. Spect. Anal. **30**(08), 2088–2092 (2010)
12. G.F. Pan, Research on the modeling method of water quality total nitrogen spectroscopic detection. Jiangnan University, 2014.
13. Y.J. Wang, L.K. Yang, Y.T. Wang, Appl. Mech. Mater. **239–240**, 1395–1398 (2012)
14. T. Houska, D. Kraus, R. Kiese et al., Biogeosciences **14**(14), 1–28 (2017)
15. Q.Q. Li, G.W. Li, J.X. Zhang, Spectrochimica acta. Part A Mol. Biomol. Spectrosc. **219**, 274–280 (2019)
16. Z. J. Shi, K. Li, J. Nat. Sci. Ed. **2**, 302–304 (2008)
17. A. Mouazen, W. Saeys, J. Xing et al., J. Near Infrared Spectrosc. **13**(1), 87 (2005)
18. B. Kuang, A.M. Mouazen et al., Biosys. Eng. **114**(3), 249–258 (2013)
19. W.Y. Lu, X.R. Zhu, S.C. Yao et al., Infrared Laser Eng. **47**(07), 155–160 (2018)
20. D. Hao, G.Y. Chen, W. Wei et al., Infrared Phys. Technol. **101**, 156–161 (2019)
21. J. Yang, J. Huang, K.E. Yuan, Infrared Laser Eng. **48**, 5 (2019)