Optimization of SGSF parameters in DOAS measurement using supercontinuum source

HE XU[a], CHENGHAI ZHAO[b], WUMING WU[a,c,d], SHENGPING CHEN[a,c,d,*], ZONGFU JIANG[a,c,d]
[a]*College of Advanced Interdisciplinary studies, National University of Defense Technology, Changsha 410073, China*
[b]*College of Science, Central South University of Forestry and Technology, Changsha Hunan 410004, P.R China*
[c]*State Key Laboratory of Pulsed Power Laser Technology, Changsha, Hunan 410073, China*
[d]*Hunan Provincial Key Laboratory of High Energy Laser Technology, Changsha, Hunan 410073, China*

The average concentrations of water vapor (H_2O) and carbon dioxide (CO_2) in the near-surface atmosphere are measured by adopting a differential optical absorption spectroscopy (DOAS) system, which uses a fiber supercontinuum source (SC) with the method of Savitzky-Golay smoothing filter (SGSF). Combining the SGSF working mechanism and the absorption features of H_2O and CO_2 together, a higher measured precision can be achieved by optimizing the SGSF parameters (window numbers and ranks) with a nearly linear relationship. Furthermore, it is indicated that a higher rank and a lower window number are more precise for measuring the gas with strong absorption. Finally, on the basis of the area rule, the concentrations of CO_2 and H_2O are calculated as 479.3 ppm and 7425.4 ppm separately, which are consistent with the detected results with the relative error of 0.46% and 0.14%.

(Received September 10, 2018; accepted April 8, 2019)

Keywords: Supercontinuum, DOAS, Measurement, Savitzky-Golay smoothing filter (SGSF)

1. Introduction

Based on the gas absorption spectroscopy, DOAS is an efficient measurement instrument for atmospheric trace gases [1, 2]. Generally, the DOAS system is composed of three parts including radiation source, target gas constituents and receiving unit [3]. Typical DOAS system employs multiple lasers with fixed or tunable wavelengths as the radiation source, such as the mid-wave infrared tunable laser [4, 5], to provide abundant spectral absorption features. The system with such radiation source was successful in measuring individual concentrations of several species [6], but failed to measure in complicated gas mixtures with overlapping spectral features due to the limited wavelength range [7, 8]. Hence, in such systems, the laser sources containing multiple wavelengths is more versatile than normal monochromatic lasers [9, 10].

The supercontinuum source (SC) not only has a sufficiently broad spectrum ranging from 400 to 2400 nm which covers the CO_2 and H_2O absorption lines, but also has high laser beam quality, such as narrow beam divergence and stable output [11-13]. By applying the SC, the DOAS system was able to observe the differential absorptions of variable atmospheric species between the ground and several kilometers in altitude simultaneously [14, 15]. Several researchers had carried out some preliminary work on applying the SC for gas measurement by means of DOAS. In the DOAS measurement of indoor CO_2 at around 2000 nm, a nanosecond white light continuum was used with a spectral resolution of 8 nm [16]. Later, another similar measurement system was set up, which is utilized to discuss H_2O and CO_2 concentration changes in daytime and nighttime during a 10-day period with a better resolution of 0.046 nm [17]. With the supercontinuum absorption spectroscopy (SAS) technique, the concentration of outdoor oxygen was measured over a 540-m path [18]. Absorption spectra of methane (CH_4), acetylene (C_2H_2), and ethylene (C_2H_4) were measured using a SC at various pressures and concentration [19]. Recently, the photo-acoustic spectroscopy technique was combined with the direct absorption spectroscopy technique to analyze the absorption of C_2H_2 [20]. Noteworthy is that, none of them had paid attention on decreasing the measuring error in data processing.

The spectral data always contains noise and disturbance signals (i.e., overlap effect), so it is significant to perform data preprocessing to obtain highly precise and accurate data [21]. As a result, many experimental schemes or techniques had been proposed for sensitivity improvement and resolution enhancement [22]. Among various filter techniques, wavelet transform (WT) is a powerful signal denoising technique [23], but this method depends on more parameters, for example, mother wavelet type, thresholding method, threshold estimation, and

decomposition level. Recently, the Savitzky–Golay smoothing filter (SGSF) seems to be especially attractive since both the smoothed signal and the derivatives can be calculated in a single step, and only two parameters (window numbers and ranks) must be set. Researches had shown the effective promotion of estimation precision by optimizing the parameters of the SGSF [24, 25]. The significance of filter shape for reducing information losses in measuring field was also proposed by Alexandre Dobroc et al [26]. Different from the previous work, we emphasize on the analysis of algorithms building process especially the parameters setting methods in detail. Optimum choice of SGSF parameters are discussed and the trend for parameters setting can be predicted based on the absorption characteristics of the gas. To the best of our knowledge, this is the first case to optimize the SGSF parameters in measurement using the SC in a DOAS system. Comparing with the previous indoor measurement results [16, 17], a better accuracy (0.46% for H_2O and 0.14% for CO_2) is achieved with a shorter optical length (200 m) and a more compact system (50 m).

2. Experimental setup

Fig. 1 shows the schematic experimental setup for the DOAS measurement of atmospheric CO_2 in the laboratory. The indoor collimated optical path length reaches about 200 m which includes three reflections by three aluminized mirrors. After targeted directly by the last retro-reflector to the receiving Cassegrain telescope, the beam is transferred to a parabolic reflector for focusing. Then it is accepted by a multimode optic fiber at the focal point into a sensitive spectrograph to get the absorption spectra. On account of the SC coverage of visible light with an average power of 700 mW, the alignment process is easy especially in the Cassegrain telescope and the parabolic collimator parts. In the optimized alignment, the spectrum data containing absorption bands of the two gas mixtures is obtained in a sampling increment of 0.024 nm.

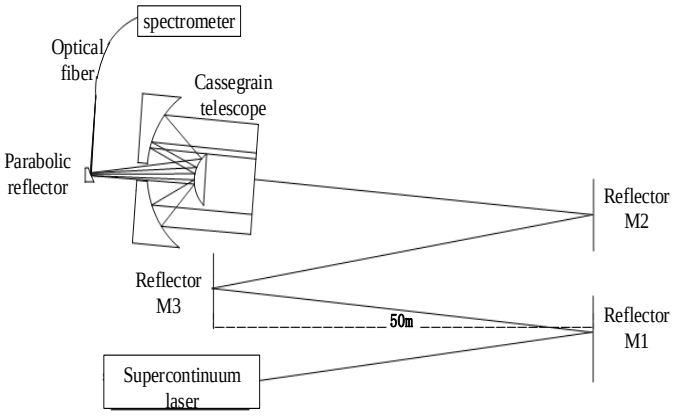

Fig. 1. Experimental setup for the DOAS measurement

3. Data analysis algorithms

Similar as standard DOAS analysis, the outline of the transmission process can be expressed as the Lambert–Beer law:

$$I(\lambda) = I_0(\lambda)\exp(-L\sigma(\lambda)c) \tag{1}$$

where $I_0(\lambda)$ is the initial illuminant intensity at wavelength λ, $I(\lambda)$ is the light intensity observed at distance L, c is the molecular number density and $\sigma(\lambda)$ is the absorption cross section of the medium molecules. $I(\lambda)$ is demonstrated in Fig. 2.

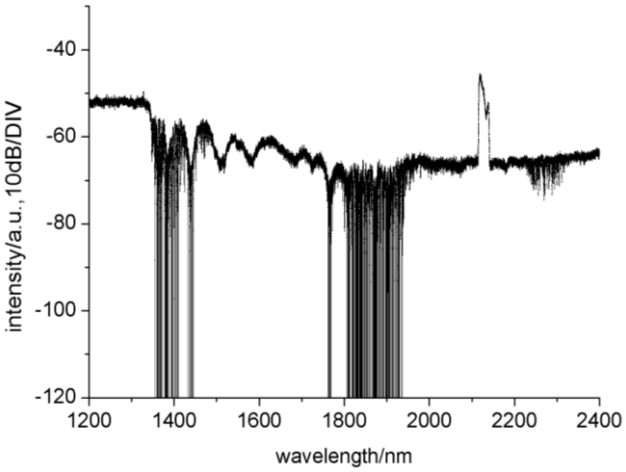

Fig. 2. The receiving light intensity at 200 m under 9.8°C, indoor, relative humidity of 83%, 1 ATM

In order to get the differential absorption cross sections $\sigma'(\lambda)$ of the target gas, which varies rapidly with wavelength, the slow varying function $\sigma_0(\lambda)$ is separated from the following equation:

$$\sigma(\lambda) = \sigma'(\lambda) + \sigma_0(\lambda) \tag{2}$$

Firstly, in order to get $\sigma'(\lambda)$, the statistics of standard gas absorption cross sections are extracted from HITRAN database to get a simulated curve. To match with the experimental data sampling, the HITRAN parameters are set as: temperature 9.8 °C, relative humidity of 83%, 1 ATM, wavenumber range 4166-8333 cm^{-1} and step 0.024 cm^{-1}. Secondly, resolution of the simulated curve is adjusted to match the spectrum resolution by a cubic spline interpolation. Then the Savitzky-Golay smoothing filter (SGSF) is used to get the smoothing de-noising trend (SDT) of standard H_2O absorption cross section. Finally, $\sigma'(\lambda)$ is educed by subtracting the standard H_2O absorption

cross section and its smoothing de-noising trend.

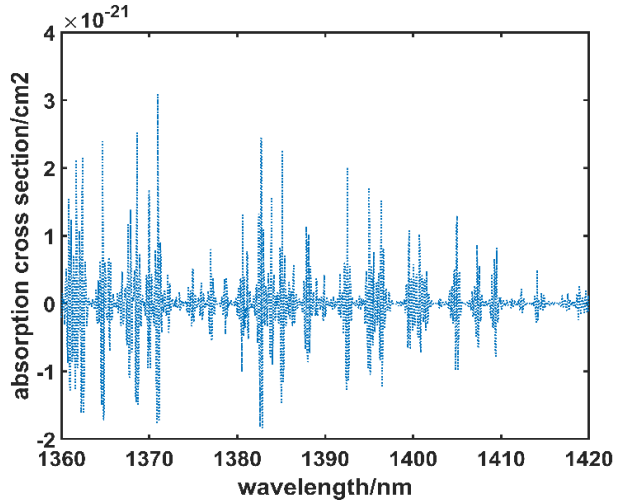

Fig. 3. The water vapor differential absorption cross section $\sigma'(\lambda)$ using information from HITRAN database under 9.8 ℃, relative humidity of 83%, 1 ATM

$I_N(\lambda)$ is the fifth polynomial of average background noise. Under optimized alignment without the SC, the background noise is collected for three times to calculate its average value. A smoother background noise line is fitted by using a fifth polynomial to reject sharp spines.

$I_0'(\lambda)$ is the SDT of the receiving light intensity simulated by the SGSF with appropriate parameters. $I_0'(\lambda)$ contains influences resulting from the optical elements, the swinging of light source and spectral distribution of detector.

$$OD'(\lambda) = \ln \frac{I_0'(\lambda) - I_N(\lambda)}{I(\lambda) - I_N(\lambda)} = L\sigma'(\lambda)c \tag{3}$$

The differential optical density ($OD'(\lambda)$) stands for the corresponding optical thickness. First, the slowly varying part of the receiving light intensity is used to divide the receiving light intensity. Then, $OD'(\lambda)$ (seen in Fig. 4) is obtained by taking the logarithm of the answer. It is not necessary to quantify $I_0(\lambda)$ after calculating $OD'(\lambda)$ using equation (3). Also, $OD'(\lambda)$ in wave bands 1360-1420 nm and 1880-1930 nm are calculated by MATLAB.

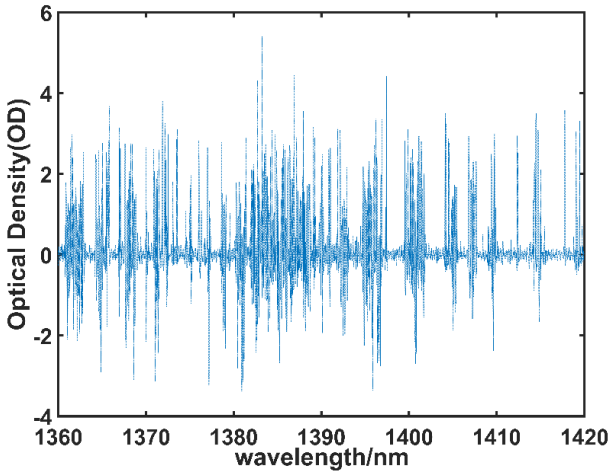

Fig. 4. The differential optical density of water vapor calculated as equation (3) under 9.8 °C, relative humidity of 83% (7421.9 ppm)

According to equation (3), there exists a linear relationship between $OD'(\lambda)$ and $\sigma'(\lambda)$. When transmitted in complicated gas mixtures, it is obvious that the total absorption of gas species ($OD'(\lambda)$) in certain waveband has linearity with the integration of $OD'(\lambda)$. Equation (4) is obtained by the integration of equation (3).

$$\int_{\lambda i}^{\lambda_h} OD'(\lambda)d\lambda = c_i L \int_{\lambda i}^{\lambda_h} \sigma_i'(\lambda)d\lambda \tag{4}$$

The atmospheric H_2O concentration is derived by average the results calculated in 1360-1420 nm (7427.6 ppm) and 1880-1930 nm (7423.2 ppm). The over-lapping absorption of H_2O is calculated by multiplying H_2O concentration with summation of $\sigma'(\lambda)$. Then, using equation (4), proportion of CO_2 is acquired by deducting the over-lapping absorption of H_2O. The ultimate results of the average H_2O concentration is 7425.4 ppm, while the hygrometer indicator is 7422 ppm (0.46%). The concentration of CO_2 comes out to be 479.3 ppm with the detected CO_2 concentration 480 ppm (0.14%).

However, there are discrepancies between the evaluation results with external data of both H_2O and CO_2. One reason is that the superposition of the first-order diffraction of light at wavelength λ and the second-order diffraction of light at λ/2. Since the supercontinuum spectrum ranges from 400 nm to 2400 nm, during the absorption process of H_2O and CO_2 in 1200-2400 nm, other absorptions in 600-1200 nm are also included [27]. Except for the superposition, the parameters setting also has a great influence on the computation.

4. Influence of parameters setting

While varying the rank from 0 to 15 and window numbers from 1 to 40, the measuring error of H_2O and CO_2 are shown in Fig. 5 and Fig. 6. It is noted that the changes of SGSF parameters (rank and window numbers) have a greater influence on calculating H_2O concentration (0%-300% error) than CO_2 concentration (0%-40% error). Generally, higher rank and window numbers are combined with smaller errors than lower rank and window numbers. Apparently, there is a best choice of parameters in the SGSF simulation process. If the rank (y) and windows number (x) fit the linear relationship illustrated as the red lines in Fig. 5 and Fig. 6 respectively, the measuring errors are the lowest. A higher rank and lower window number are more appropriate for H_2O measurement while a relatively lower rank and higher window number for measuring CO_2. The reasons for the parameters setting can be explained by how the SGSF works and the absorption features of the two gases.

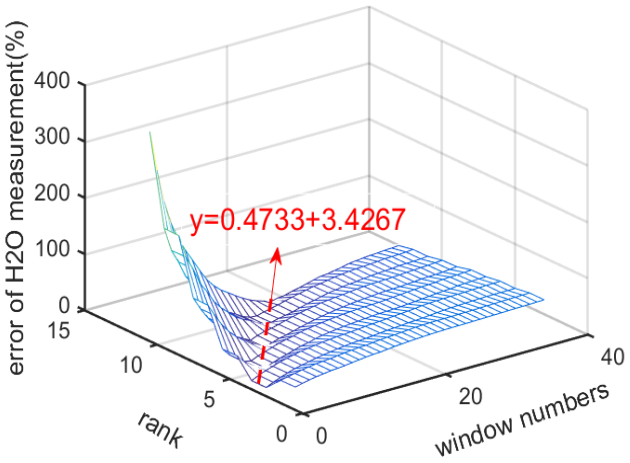

Fig. 5. H_2O measuring error in different parameters

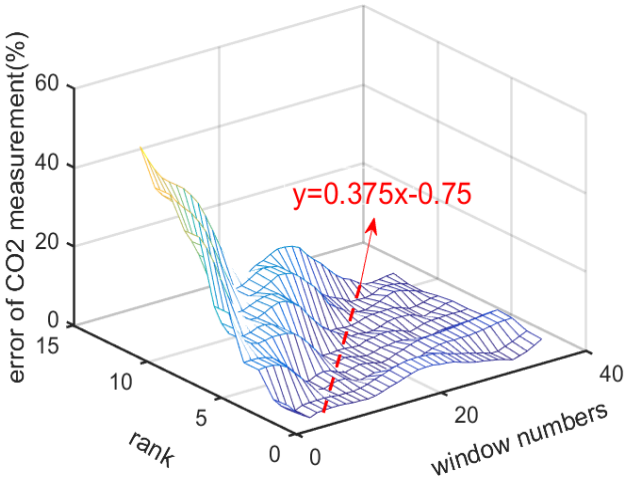

Fig. 6. CO_2 measuring error in different parameters

The SGSF works as follows [28]: the window number and the rank are n (=2m+1) and k respectively, then the number of points are x=(-m, -m+1,…,0,…, m-1, m) and the fitting equation for one point is $y = a_0 + a_1 x + a_2 x^2 + \cdots a_{k-1} x^{k-1}$, so for all the (2m+1) points

$$\begin{pmatrix} y_{-m} \\ y_{-m-1} \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} 1 & -m & \cdots & (-\mathrm{m})^{k-1} \\ 1 & -m+1 & \cdots & (-\mathrm{m}+1)^{k-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & m & \cdots & -\mathrm{m}^{k-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{k-1} \end{pmatrix} + \begin{pmatrix} e_{-m} \\ e_{-m+1} \\ \vdots \\ e_m \end{pmatrix} \quad (5)$$

$Y = \begin{pmatrix} y_{-m} & y_{-m-1} & \cdots & y_m \end{pmatrix}^T$ is the filtered value.

The absorption features of H_2O and CO_2 can be observed from Fig. 7 and Fig. 8. Obviously, less spikes and oscillations with smaller scales consists in waveband 2000-2030 nm (absorption waveband of CO_2) compared with the spectra in 1360-1420 nm (absorption waveband of H_2O). Therefore, the H_2O absorption is higher and more complicated than CO_2.

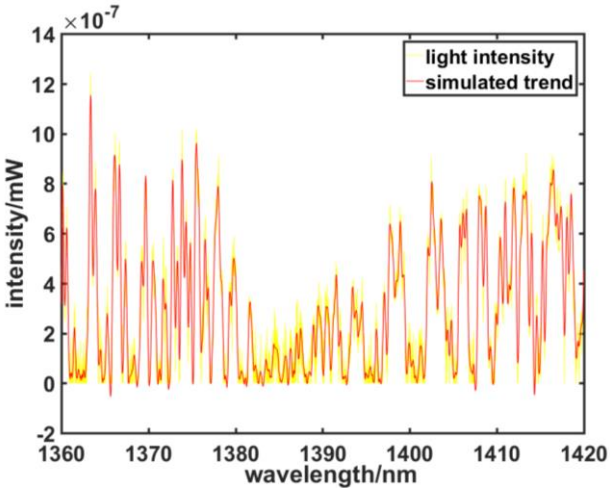

Fig. 7. Comparison between light intensity at 200 m and its smoothing trend in 1360-1420 nm

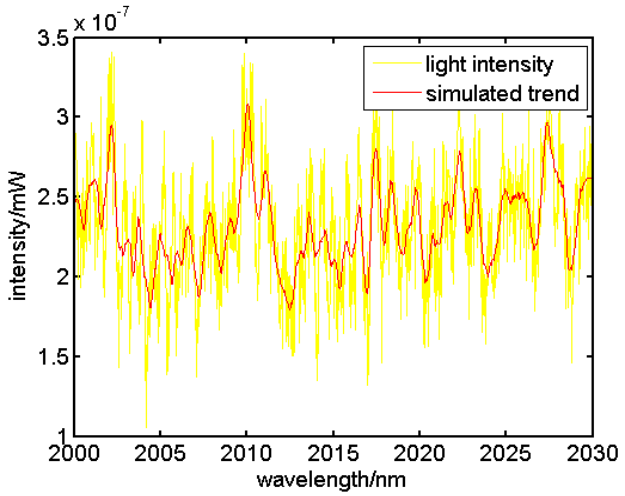

Fig. 8. Comparison between light intensity at 200 m and its smoothing trend in 2000-2030 nm

Combining the SGSF working mechanism with the differences between two gases, the phenomenon for SGSF parameters above can be explained. For CO_2 calculation, as it varies slightly in absorption lines, more sampling points are needed but a low-order polynomial is enough to characterize the trend, indicating that a lower rank and a higher window number are more appropriate for its measurement. Meanwhile, the SGSF parameters setting would lead to a smaller error on the calculation of such weak absorption gases. As for H_2O, the strong absorption and drastic change require a high-order polynomial. So the proper parameters for H_2O are a higher rank and lower window number. As for gases like H_2O with a strong absorption, the setting for SGSF parameters would lead to a larger error on its calculation. In addition, no matter for strong or weak absorption gases, the optimized rank and window numbers meet a nearly linear relationship.

5. Conclusions

In order to improve the measured resolution of water vapor and CO_2 concentrations, the DOAS system is used together with SGSF method which provides an alternative way to SAS combined with maximum likelihood estimation (MLE) estimation algorithm in gas measuring field [18]. The results show that among all the DOAS system utilizing the SC, a better accuracy in the atmospheric H_2O and CO_2 measurement is achieved with the most compact system and shortest optical length [16, 17]. The SGSF parameters setting is discussed in detail in the algorithms building process with interesting phenomena, which results from the SGSF working mechanism and the absorption features of gases. To reduce the measuring errors, the rank and windows number are set to fit a nearly linear relationship. Furthermore, it is indicated that a higher rank and a lower window number are more precise for measuring gases with more obvious absorptions. It should be noted that in order to concentrate on the measured resolution of water vapor and CO_2 concentrations, only one case is conducted in the present paper; however, more other circumstances should be considered in the future work.

Acknowledgments

The work was supported by the National Natural Science Foundation of China under Grants 11504424.

References

[1] M. Liebl, Differential optical absorption spectroscopy (DOAS), 2012.
[2] J. Chong et al., SPIE Remote Sensing, Proc. SPIE **10001**, 7 (2016).
[3] P. P. Geiko, S. S. Smirnov, I. V. Samokhvalov, 20th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, Proc. SPIE **9292**, 8 (2014).
[4] P. Weibring et al., Applied Physics B **79**(4), 525 (2004).
[5] W. C. Perkins et al., SPIE BiOS, Proc. SPIE **8926**, 5 (2014).
[6] M. Xiao et al., Spectroscopic characterization. Spectrochimica Acta Part A Molecular & Biomolecular Spectroscopy **125**(6), 120 (2014).
[7] K. I. Arshinov et al., 15th Symposium on High-Resolution Molecular Spectroscopy, Proc. SPIE **6580**, 6 (2006).
[8] U. N. Singh et al., SPIE Asia-Pacific Remote Sensing, Proc. SPIE **9879**, 10 (2016).
[9] U. Minoni et al., Optics & Laser Technology **49**(7), 91 (2013).
[10] H. J. Patrick et al., Proc. SPIE - The International Society for Optical Engineering **8495**, 84950K (2012).
[11] K. Yin, et al., Journal of Lightwave Technology **35**(20), 4535 (2017).
[12] R. Song et al., SPIE Defense + Security, Proc. SPIE **10192**, 7 (2017).
[13] V. Choudhury et al., SPIE LASE, Proc. SPIE **10513**, 6 (2018).
[14] J. Hult, R. S. Watt, C. F. Kaminski, Optics Express **15**(18), 11385 (2007).
[15] D. M. Brown et al., Optics Express **16**(12), 8457 (2008).
[16] T. Somekawa et al., Opt. Lett. **36**(24), 4782 (2011).
[17] H. Saito et al., Opt. Lett. **40**(11), 2568-71 (2015).
[18] D. M. Brown et al., Journal of Applied Remote Sensing **8**(1), 083557 (2014).
[19] J. Yoo et al., Applied Spectroscopy **70**(6), 1063 (2016).
[20] R. Selvaraj, N. J. Vasa, S. M. S. Nagendra, Lasers and Electro-Optics Europe & European Quantum

Electronics Conference **1**, 1-1 (2017).
[21] P. J. Abraham et al., Spectrochimica Acta Part A, Molecular and Biomolecular **71**, 355-67 (2008).
[22] P. Fjodorow et al., Applied Physics B **120**(4), 667 (2015).
[23] J. Li, B. Yu, H. Fischer, Applied Spectroscopy **69**(4), 496 (2015).
[24] P. I. Stavroulakis et al., SPIE Optical Engineering + Applications, Proc. SPIE **8842**, 16 (2013).
[25] H. Deng et al., International Symposium on Photonics and Optics, Proc. SPIE **9656**, 6 (2015).
[26] A. Dobroc, N. Cezard, Applied Optics **51**(35), 8470 (2012).
[27] T. Jin, J. Zhou, P. T. Lin, SPIE Defense + Security Proc. SPIE **10629**, 4 (2018).
[28] Y. Q. Chen, Savitzky-Golay Smoothing Filter 2010: Betascript Publishing.

*Corresponding author: chespn@163.com