 doi:10.20944/preprints201812.0331.v1

Article

Extreme Learning Machine for Robustness Enhancement of Gas Detection Based on Tunable Diode Laser Absorption Spectroscopy

WENHAI JI [1], LI ZHONG [1], YING MA [1], DI SONG [1], XIAOCUI LV [1], CHUANTAO ZHENG [2], GUOLIN LI [1,*]

[1] College of Information and Control Engineering, China University of Petroleum, Qingdao 266580, China
[2] State Key Laboratory on Integrated Optoelectronics, College of Electronic Science and Engineering, Jilin University, Changchun, 130012, China
* Correspondence: liguolin@upc.edu.cn; Tel.: +86-532-86980953

Abstract: In this work, a tailored extreme learning machine (ELM) algorithm to enhance the overall robustness of gas analyzer based on the tunable diode laser absorption spectroscopy (TDLAS) method is presented. The ELM model is tailored through activation function selection, input weight and bias searching, and cross validation method to address the analyzer robustness issues for industrial process analysis field application. The two particular issues are the inaccurate gas concentration measurement caused by the process gas background components variation, and the inaccurate spectra shift calculation caused by spectral interference. By using our algorithm, the concentration error is reduced by one order of magnitude over a much larger stream pressure and component range compared with that obtained by classical least square (CLS) fitting methods based on reference curves. Additionally, it is shown that with our algorithm, the wavelength shift accuracy is improved to less than 1 count over 1000 counts spectra length. In order to test the viability of our algorithm, a trace ethylene (C_2H_4) TDLAS analyzer with coexisting methane was implemented, and its experimental measurements support analyzing robustness enhancement effect.

Keywords: gas analyzers; optical sensors; TDLAS; extreme learning machine

1. Introduction

Absorption spectroscopy is a versatile technique to obtain fundamental physics parameters with great precision. Tunable diode laser absorption spectroscopy (TDLAS) technology has gained popular acceptance for trace gas measurement in various industries due to rapid development of optoelectronics technology. Because of its high sensitivity and fast response time, it is widely used in coal mine safety [1], combustion measurement [2], petrochemical process analysis [3], dissolved gas detection [4], gas emission and air quality monitoring [5] etc.

But there are challenges to maintain the analyzing robustness and accuracy for long term onsite industrial applications. The conventional concentration measurement method is based on calibration model using reference spectra [6, 7]. The common issues are spectral deformation and spectral interference caused by process background and operating condition variation. These factors are entangled in real time measurement due to the dynamic nature of industrial process. The analyzing system deviates from the original calibration and its performance degrades over time. To address these issues, linear regression algorithms such as classical least square fitting (CLS, also known as multiple linear regression -MLR) method [6], partial least square fitting (PLS) method [7] are commonly adopted in TDLAS spectra analysis. CLS algorithm is also named as reference curve

 doi:10.20944/preprints201812.0331.v1

method. PLS is an integration of regression and principle component analysis. The models are good for certain tolerance of background change.

But when the spectral interference or deformation becomes stronger, it is hard to achieve satisfactory performance with the linear regression algorithms. Then nonlinear algorithms such as artificial neural networking (ANN) [8], supporting vector machine (SVM) [9] and extreme learning machine (ELM) [10-15] were introduced in spectra analysis. Among those, ELM is a novel single layer forward neuro networking (SLFN) algorithm proposed and developed by G. Huang [16-20]. It has the relative advantages of simple structure and less computation time. ELM is capable of regression and classification [21-22]. For the classification application, ELM was used in coal classification [10], dynamic fresh fruit detection [11], and traffic road signs identification [22] etc. The regression ELM for quantitative calculation was applied in element analysis [12], gas concentration retrieval [13], stellar star abundance calculation [14], diesel fuel and blend oil analysis [15], and food extrusion processing [23] etc. ELM spectra analysis [10-15] was used in micro-NIR, Fourier transform infrared (FTIR) spectroscopy, laser induced breakdown spectroscopy (LIBS) area. Its application in TDLAS spectra analysis for industrial process analysis was not implemented so far.

2. Experiment

The process of TDLAS analyzer gas concentration measurement is shown in Figure. 1. The spectra generation process is illustrated in Figure. 1 (a). The second harmonic component of laser absorption was obtained through laser current modulation and demodulator extraction. The procedure of calibration model training and verification is illustrated in Figure. 1(b) and (c). However, in the practical model verification, there are challenges due to the harsh operating environment. Particularly, we will deal with two issues. First issue is the accurate gas concentration prediction caused by background components and gas pressure variation during process measurement. Second issue is the accurate spectra shift prediction with strong spectral interference. The direct impact is the degradation of prediction accuracy, shown by the error band width d change from ideal case to practical case in Figure. 1(c). ELM was proposed to address the challenges. With ELM, we aim to calculate both the gas concentration in dynamical process background and the actual spectra shift with the interfering spectral structures.

We choose trace ethylene (C_2H_4) measurement as the application example. C_2H_4 is an important indicator gas which needs real time monitoring at ppmv (part per million by volume, default as ppm) level in many areas. For example, in the biomedical area, analyzing C_2H_4 in breath air can be used to study lipid peroxidation of human body cell, so that related deases can be diagnosed at early stages. Another example is monitoring coal spontaneous combustion which is a major risk in the coal mining wells and storages [24-25]. Trace level C_2H_4 is produced in early stages and the concentration increases as the process continues. CH_4 is a common coexisting component and also brings spectra interference. So, accurate measurement of C_2H_4 concentration at trace level is difficult.

Accordingly, 0-200ppm range C_2H_4 measurement experiment was designed with 0-2000ppm CH_4 presence and in dynamic environment. The system schematic is shown in Figure. 2. A Cortex-A5 chip manufactured by Atmel is used as the master micro controller unit (MCU). The A5 MCU controls a commercial module PCI-FPGA-1A manufactured by Port City Instruments (PCI). Its main functions are laser wavelength scan, modulation, photoelectric current amplification and demodulation. A sawtooth shape driving current is applied to the distributed feedback (DFB) laser to scan the wavelength. The tuning range is around 0.45 nm which covers two absorption peaks of the C_2H_4 at 1626.35nm and 1626.53nm respectively. The laser temperature is stabilized and adjusted by a WTC3293 module (TEC, temperature controller). The scan period is 102ms and the maximum out of fiber power is 3.3mW. The laser driving current is superimposed with a 31.4 KHz sin wave AC signal (1f) for amplitude modulation. The transmitted light is collected by a PIN InGaAs detector. The photon current is directed sequentially through I/V convertor, preamplifier, band pass filter, and phase sensitive lock-in amplifier (demodulator). The second harmonic component denoted as 2f is generated by using a frequency doubled reference signal (2f Ref) in demodulator. The modulation signal (1f) and demodulation reference signal (2f Ref) are generated through a digital signal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

synthesizer (DSS). The optical power signal during the scan denoted as DC is obtained through a low pass filter. DC and 2f, together with gas status parameter temperature T and pressure P are feed to the A5 processor through an analog to digital convertor (ADC).

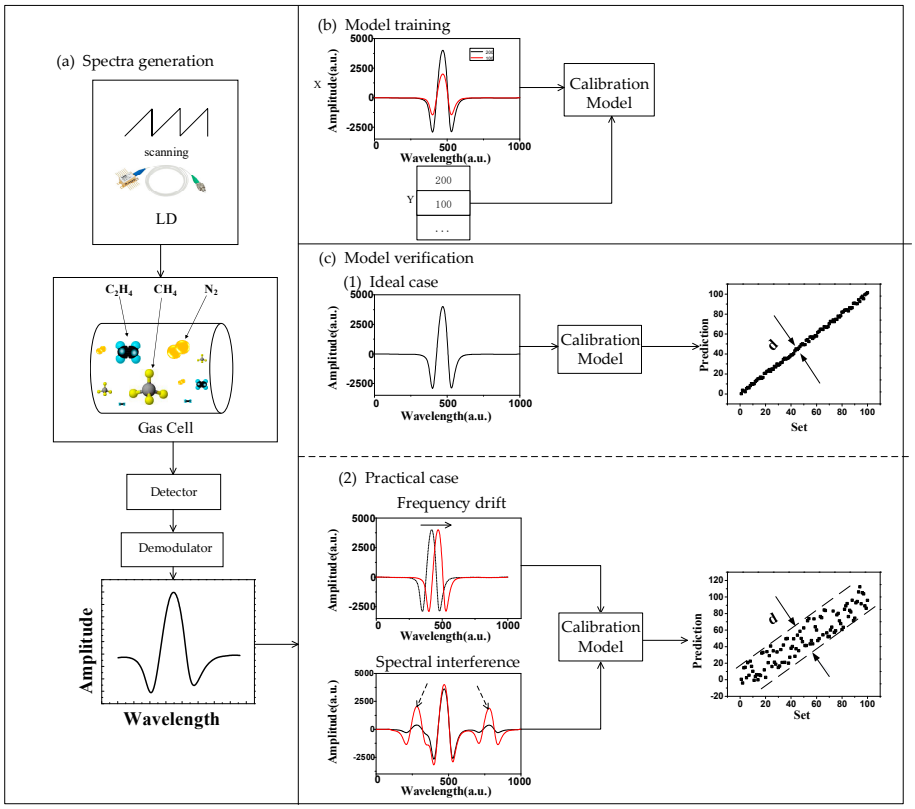

Figure 1. The process of TDLAS analyzer gas concentration measurement. (a) 2f spectra generation; (b) Model training with training set; (c) Model verification with collected spectra.

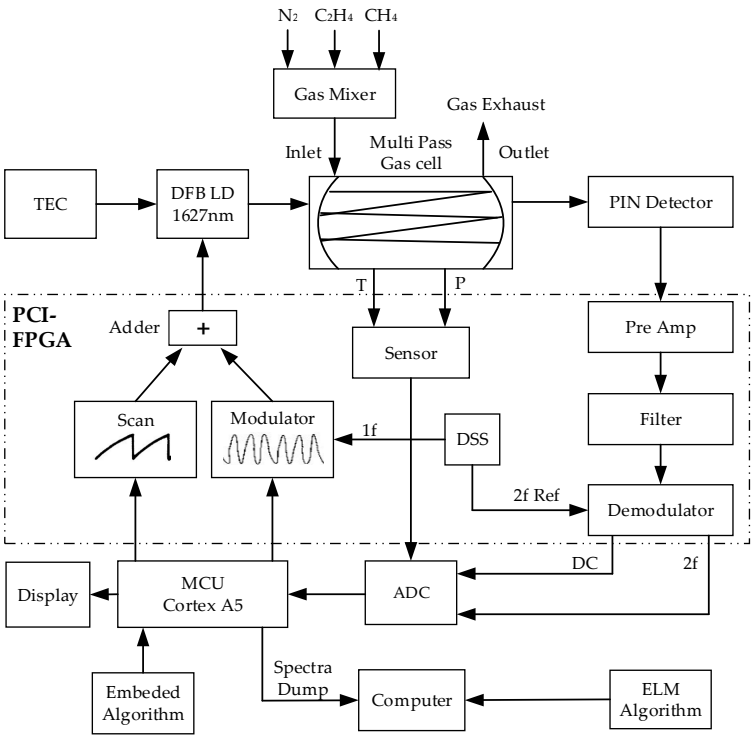

Figure 2. Schematic of C_2H_4 measuring TDLAS platform.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

The C_2H_4 absorption features were thoroughly searched in the near infrared region [24, 26]. Absorbance of 1ppm-m moisture (H_2O), carbon monoxide (CO), carbon dioxide (CO_2), CH_4 and C_2H_4 in 1626-1627nm wavelength region is depicted in Figure 3. (a). To demonstrate the clear features of CH_4 spectra, its absorbance is plotted on right axis whose scale is 10 times of left axis. At 25℃, atmosphere pressure and 1ppm-m concertation level, the peak absorbance of CH_4 is one magnitude weaker than C_2H_4. The other potential components such as H_2O, CO, and CO_2 are four or more magnitudes weaker than C_2H_4. Thus, from 1626.25nm to 1626.70nm region (defined by two magenta dotted lines in Fig. 3(a)) is the ideal wavelength scan range for C_2H_4 analysis. Experimentally acquired DC and 2f spectra of 200ppm C_2H_4 and 2000ppm CH_4 balanced in N_2 are shown in Figure. 3(b) and used as the reference curve in CLS method. The spectral structure of C_2H_4 and CH_4 are overlapped. The gas stream is generated through an automatic gas mixing station with premixed 1000 ppm C_2H_4 in nitrogen (N_2), 1% CH_4 in N_2 and pure N_2 bottles.

To address the spectra analysis challenges, two experiments were designed. To prepare for analyzer calibration model, test samples should be representative and cover all possible status. Thus, the associated independent variables of the samples such as gas concentrations, pressure, laser driving current *etc.* should be randomly set. The first experiment was the process simulation test, where CH_4 was mixed in stream. The C_2H_4, CH_4 concentration, and gas pressure was randomly varied through the Excel random function in experiment design. The second experiment was spectra shift simulation test, where driving current was randomly adjusted with dynamic gas composition to mimic spectra drift effect.

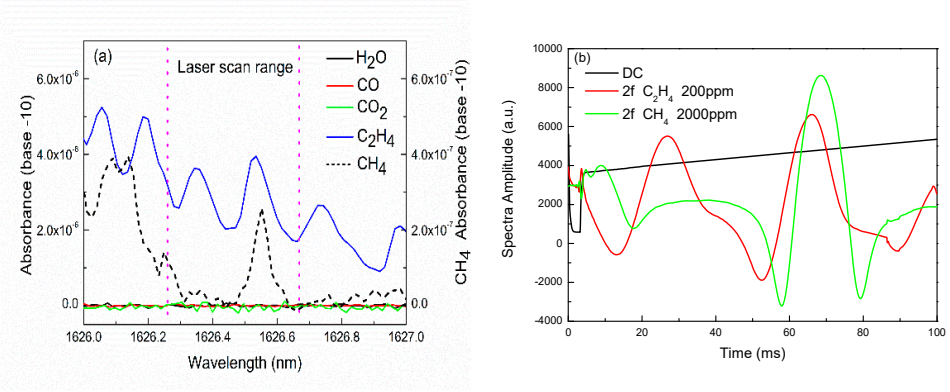

Figure 3. (a) Absorbance of 1ppm-m H_2O, CO, CO_2, CH_4 and C_2H_4 in 1626-1627nm wavelength region. (b) Experimentally acquired DC and 2f spectra of 200ppm C_2H_4 and 2000ppm CH_4.

2.1. Process Simulation Experiment

This experiment was to simulate the industrial process in a controlled scenario, where the C_2H_4 concentration, CH_4 concentration and gas pressure varied during the measurement. Fifty mixing gas streams were generated with C_2H_4 up to 200 ppm, CH_4 up to 2000 ppm and balancing nitrogen. The pressure was controlled in 900 to 1100 mbar range. Gas cell was maintained at room temperature. In order to get representative test sample, the concentrations and pressure were randomly set. The 50 corresponding spectra are plotted in Figure. 4 (a). For an isolated spectra feature, the gas concentration is proportional to the spectra peak height, which is the simple case of Beer-Lambert law. But due to CH_4 interference, accurate measurement can't be obtained by direct counting the C_2H_4 spectra peak height.

2.2. Spectral Drift Simulation Experiment

A commercial analyzer can be used for 5-10 years. During its lifetime, spectra deformations are unavoidable. Many factors lead to deformation, including the complex operating environment, the degradation and aging of laser, detectors, and electronics components. Among all the deformations,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

spectra drift is the most common type and produces adverse impact on the calibration model accuracy.

In this experiment, the gas cell pressure was fixed at 1000 mbar. The gas concentration was also randomly set. The starting current was randomly set in between 59.8 mA to 60.2 mA. The variation range was 0.4 mA and the peak shift range was around 16 counts over 1000 counts spectra range. The current scanning range was fixed at 24 mA. Thus, there was no stretching or shrinking effect. The collected spectra are shown in Figure.4 (b). For an isolated spectra feature, the spectra shift can be calculated by peak tracking based on peak position difference between current spectrum and target spectrum. But it is unlikely to get spectra shift by tracking any spectra peak for the overlapping spectra structures in Figure. 4(b).

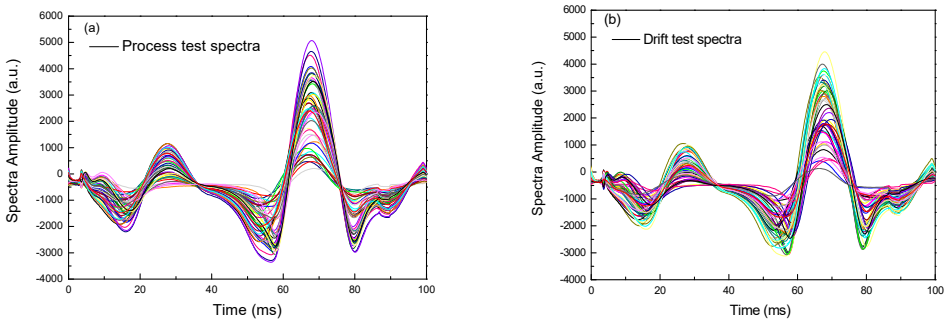

Figure 4. Spectra collected in the process simulation test (a) and in spectra drift simulation test (b).

3. ELM Algorithm Theory

ELM was originally developed for the single-hidden-layer feedforward neural networks and then extended to the "generalized" SLFNs whose hidden layer need not be tuned [16–20]. Essentially, ELM originally proposes to apply random computational nodes in the hidden layer, which are independent of the training data. Different from traditional learning algorithms for a neural type of SLFNs [16], ELM aims to reach not only the smallest training error but also the smallest norm of output weights. ELM is identified with better training speed and generalization ability than BP neuro networking and SVM. The ELM algorithm structure is shown in Figure. 5.

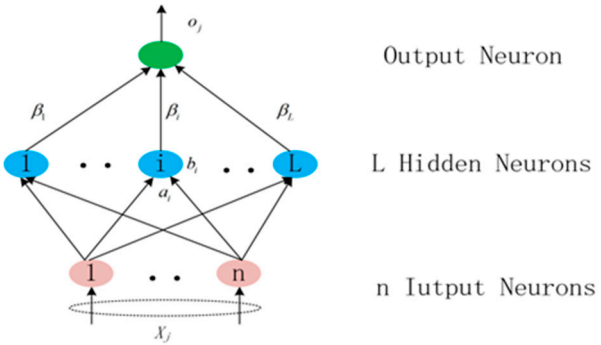

Figure 5. Structure of ELM algorithm.

We target to improve the prediction accuracy of C_2H_4 concentration or spectra shift amount. They are assigned as model output o. The prediction is based on the collected spectra. Other variables such as gas temperature T and pressure P may also be helpful to improve prediction accuracy. They are assigned as model input X.

The model output o is expressed as,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

$$\sum_{i=1}^{L}\beta_i g(W_i \cdot X_j + b_i) = o_j, j = 1,\cdots,N, \quad (1)$$

where β_i is the output weight, $W_i = [\omega_{i1}, \omega_{i2}, \cdots, \omega_{in}]^T$ is input weight, X_j is the j-*th* input of spectra sample X, $g(\cdots)$ is the activation function, $W_i \cdot X_j$ is the inner product, b_i is the bias of the neuron, o_j is the output with respect X_j, N is the number of the sample, L is the number of hidden layer nodes, n is the input layer neuron number, the output layer neuron number is 1.

The goal of ELM model training is to find β_i, W_i and b_i to match the real output t with model prediction o,

$$\sum_{j=1}^{N} \| o_j - t_j \| = 0, \quad (2)$$

Let

$$\sum_{i=1}^{L}\beta_i g(W_i \cdot X_j + b_i) = t_j, j = 1,\cdots,N, \quad (3)$$

or in matrix form,

$$H\beta = T, \quad (4)$$

where,

$$\mathrm{H}=\begin{bmatrix} g(W_1 \cdot X_1 + b_1) & \cdots & g(W_L \cdot X_1 + b_L) \\ \vdots & \cdots & \vdots \\ g(W_1 \cdot X_N + b_1) & \cdots & g(W_L \cdot X_N + b_L) \end{bmatrix}_{N \times L}, \quad (5)$$

As a result, the output weight is estimated after model training as,

$$\hat{\beta} = H^{+}T = H^{+}(HH^{T})^{-1}T, \quad (6)$$

Compared with SLFNs, the major difference of ELM is that the input weight matrix W and bias matrix b of the hidden layer is randomly initialized. Only the number of hidden layer neurons L is to be optimized. That's why ELM is so distinct in its simple structure and fast learning speed. The problem of low learning efficiency and tedious parameter setting of traditional ANN algorithm is not an issue in ELM.

ELM optimization research was focused on kernel or excitation function selection [19], progressing modelling and generic algorithm [23], priori parameter initialization [27], and particle swarm optimization [28] etc. ELM model can be improved through optimizing the neuron numbers L [16]. Due to the randomness of the input weight and threshold, there exists fluctuations in the final output. Its stability and generalization ability could be further improved by generic algorithm (GA) optimization of input weight W and bias b. GA is well known for searching of global optimized parameter [23]. GA searching is done through many generations of selection, crossover and mutation operation. Since GA is well studied, its detailed theory is not presented.

The activation function is also critical to the model performance in terms of model convergence speed and predication accuracy. But there is no general rule of activation function selection. In neuro networking algorithms, the popular activation function choices are sigmoid, tanh, and ReLU [29] etc. The impact of activation function needs further study with experiment data.

To compare the performance of ELM, the conventional classical least square (CLS) fitting method is introduced with following expression,

$$Y = a_0 + a_1 R_1 + a_2 R_2, \quad (7)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

where Y is the collected spectra, R_1 and R_2 are reference spectra， a_i (i=1,2) is the corresponding regression coefficient which is used to determine the gas concentration.

4. Algorithm model evaluation and discussion

4.1. Experimental Spectra Analysis Procedures

Spectra were analyzed with the following procedures: spectra collection, spectra pre-process, spectra partition, model training, model evaluation, model optimization, as shown in Figure. 6.

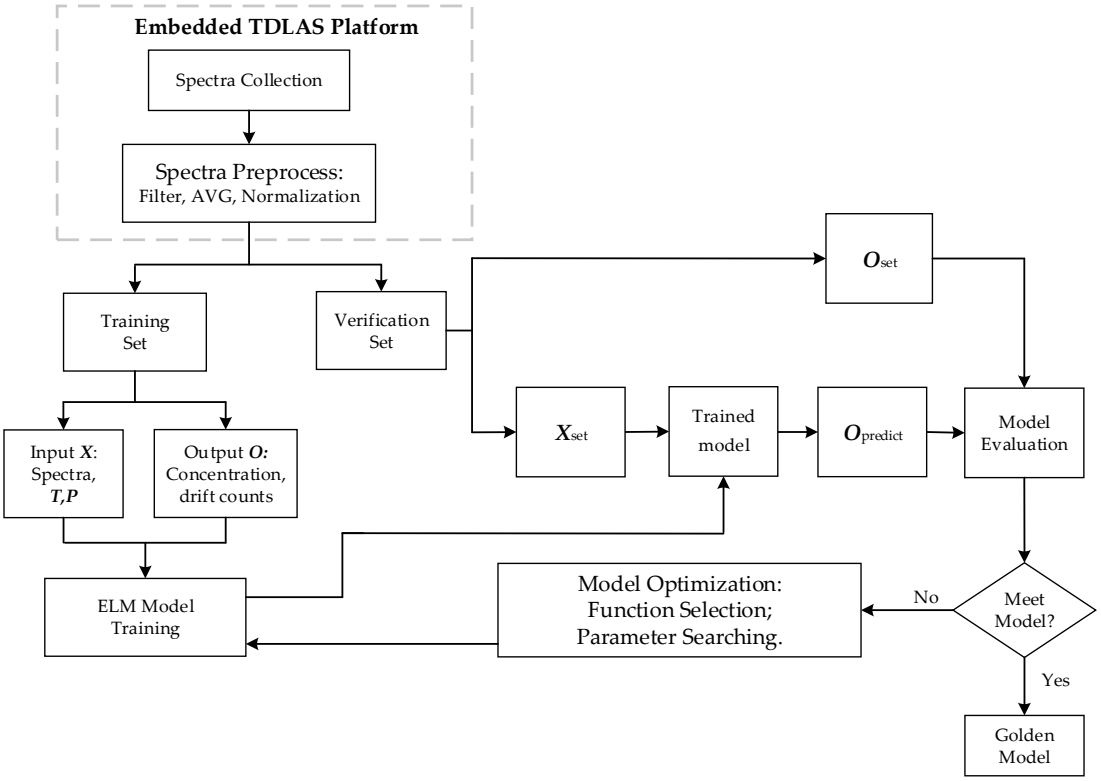

Figure 6. Spectra analysis flow chart.

Spectra collection and pre-process is the first step and done in the embedded system. The collected spectrum has 1000 data points. The spectra were averaged after 50 scans. They were further smoothed by Savitzky–Golay (S-G) polynomial filter to improve the signal to noise ratio (SNR). To cancel the effect of light power fluctuation on spectra strength, spectra were normalized by DC. The PCI-FPGA module separates the spectra into the laser off null period and scan period. The demodulation was done through software, it produces a delay in 2f spectra with respect to DC and needs further alignment. So the region from 80 to 860 was selected for analysis to exclude null period and trivial 0s from spectra alignment.

In spectra partition, they were divided into verification set and training set. A popular method is the k-fold cross validation (CV). We take 5-fold CV as an example. Spectra were first labeled by indices. They were separated into 5 subgroups per the remainder after the spectra index was divided by 5. Spectra in 4 subgroups were used for model training. Spectra in the leftover subgroup were used for verification. The 5 subgroups took turns to act as training and verification sets.

The model performance can be evaluated in many respects such as training time, computation resources, prediction accuracy etc. For the gas analysis purpose, the accuracy is the most significant factor. So, we introduce scale error,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

$$\begin{aligned} Scale\_Error_i &= Abs(\frac{o_i - t_i}{Fs}) \\ Scale\_Error &= Max(Scale\_Error_i) \end{aligned}, \quad (8)$$

where o_i and t_i are the prediction and set values of the i-*th* spectrum in verification set, Fs is the full measurement range. Depending the analyzing variables, the accuracy formula may take some modifications. The spectra shift prediction was converted to the current changes in mA.

4.2. Process Simulation Test Analysis

The 50 spectra collected in the process simulation test are used to calculate the C_2H_4 concentration prediction in dynamic backgrounds with ELM model. There are several aspects that effect the model performance, such as the function options and parameters value. It is done through the following optimization step.

First, we study the effect of activation function with sigmoid, sin, hardlim, tribas, radbas and rectified linear unit (ReLU) choices with subgroup 1 to 4 as training set and subgroup 5 as verification set. The scale errors for different activation functions are listed in Table 1. The scale error for ReLU is 0.65% which is substantially better than other choices. By comparing the analytical expression of different activation functions [29], we find that ReLU has linear expression and broad linear response input range without saturation. In this research scenario, the spectra has linear or quasi-linear response to concentration, pressure, temperature or drift variation in defined range. Thus, ReLU is more favorable to the spectra analysis compared with other activation functions than other choices. It is also faster in convergence and its behavior is closer to authentic biological neural activation function. So ReLU is selected for future ELM model.

Table 1. Model performance indicated by scale error with different activation functions.

Activation Function	Scale Error
Sigmoid	14.74%
Sin	78.90%
Hardlim	8.96%
Tribas	85.50%
Radbas	85.50%
ReLU	0.65%

Then, the GA global searching is used to find optimum hidden layer neuron number *L*, input weight *W* and bias b. In GA search setting, the group scale is 50, max generation is 25, crossover rate is 0.93, and mutation rate is 0.05. *L* is optimized at 150. After W_{best} and b_{best} are identified at generation 12, the scale error is improved from 0.65% to 0.37%.

Finally, a 5-fold cross validation method is used for the model evaluation. In the validation, each of the 5 groups take turns to act as verification set to test the model trained from the rest four groups. Scale error of each test subgroup with corresponding ELM model is listed in Table 2. It is obvious that ELM model built with subgroup 2-5 as training and subgroup 1 as verification has lowest scale error 0.27%.

The optimized ELM models are obtained after GA parameter searching, activation function trial, and 5-fold cross validation. The test error which is defined as prediction values subtract set value normalized by the full range is plotted in Figure. 7(circle). The optimum ELM model is obtained with ReLU activation function and with k=1 subgroup cross validation. The errors of training set are closer to zero because they are deeply suppressed by the activation function during the model training process in which the predication matches the set value as close as possible. The errors of 10 spectra in verification set are larger than in training set. The ELM model performance is represented by max error in the corresponding verification subgroup.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

Table 2. Scale error in each subgroup for CLS and optimized ELM models.

Subgroup	CLS	ELM
1	5.70%	**0.27%**
2	2.80%	0.40%
3	6.00%	0.46%
4	4.20%	0.57%
5	**6.30%**	0.37%

As comparison basis, CLS is introduced to obtain the C_2H_4 concentration. Its performance is represented by the max error of 50 sample spectra. In CLS regression, R1 is taken with 200ppm C_2H_4 and R2 is taken with 2000ppm CH_4. The scale error of CLS method for 50 spectra is listed in Table 2 with max error of 6.30%. The CLS test error is also plotted in Figure. 7 (triangle). From the Table 2 and Figure. 7, the error of ELM model is reduced by one order of magnitude than traditional CLS method.

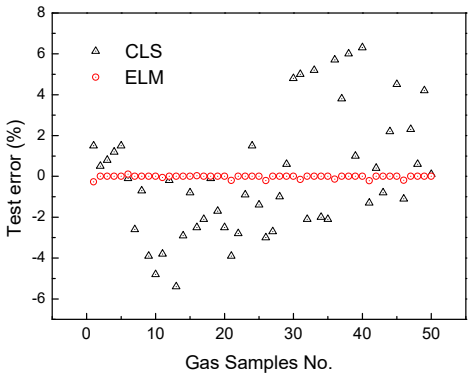

Figure 7. C_2H_4 concentration prediction errors vs gas samples for CLS method and optimized ELM model.

4.3. Spectral Drift Simulation Test Analysis

4.3.1. Spectra shift measurement with ELM model

The ELM algorithm was still utilized to obtain the actual spectral shift in the dynamic process. After GA optimization and 5-fold cross validation, the neuron number is 170 and the activation function is ReLU. In the modelling C_2H_4 concentration is replaced by spectral shift in the model. The scale error is defined as the current prediction error divided by the 16 counts variation range. The scale errors for subgroups are listed in Table 3.The best model corresponding to scale error of 4% is ELM model trained by subgroup 1,2,4,5. The spectra shift prediction error is 4%×16= 0.64 count.

In conventional peak tracking method, spectra shift is monitored the by tracking the peak with least spectral interference. So the right C_2H_4 peak is selected. The scale error of each subgroup is listed in Table 3. Due to concentration variation and residue CH_4 interference, the precision and accuracy of spectra wavelength shift calculation is much worse than ELM method. The max error is 100% and 16 counts.

Table 3. Scale error of spectra shift calculation in each subgroup for peak tracking and ELM.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 December 2018 doi:10.20944/preprints201812.0331.v1

Subgroup	Peak Tracking	ELM
1	-56%	-16%
2	-75%	-15%
3	69%	4%
4	63%	10%
5	-100%	5%

4.3.2. Prediction error after shift correction

After the spectra shift obtained through peak tracking and ELM model, the spectra can be backshifted. The purpose is to match the shifted spectra with calibration spectra so that original calibration model is still valid. The C_2H_4 concentration prediction error with and without backshift correction are calculated with previously obtained optimum ELM model and CLS method. The result is shown in Figure. 8 (a) and (b).

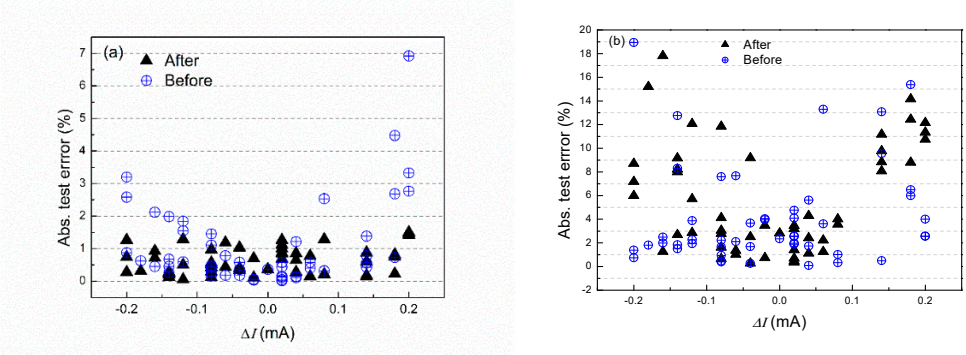

Figure 8. C_2H_4 concentration prediction abs. error vs. current shift with and w/o shift correction. (a) ELM method (b) peak tracking method.

The scale error of C_2H_4 concentration for each spectrum is plotted against ΔI ,

$$\Delta I = I_{set} - 60, \tag{9}$$

where I_{set} is the setting current, 60 mA is the base current. Before correction, the error pattern demonstrates a V shape. The explanation is that as spectra shifts more, the error increase more. After ELM correction, the error patterns demonstrate flat belt shape. The belt widths are greatly shrunk. The error is improved from 6.92% to 1.50%. For peak tracking and CLS method, the error has no obvoius improvement. This is due to the fact of very inaccurate peak shift calculation. The improvement of spectra shift correction with ELM method is demonstrated.

5. Conclusions

In this paper, ELM were employed in the TDLAS spectra analysis. Two experiments were carried out with trace C_2H_4 measurement as the application example. In the process simulation test, not only the C_2H_4 and CH_4 concentration were varied, but also the gas pressure was controlled in 900 to 1100 mbar range. For optimized ELM model, scale error of C_2H_4 concentration is obtained at 0.27% with ReLU activation function, GA optimization of input neuron number L, input weight W and bias b, and 5-fold cross validation. For traditional CLS method, the scale error is 6.30%.

In the spectra shift simulation test, ELM and peak tracking method are used to analysis the spectra shift. The prediction errors of spectra shift are less than 1 count for ELM and 16 counts for peak tracking. Then the spectra were back shifted to match with the calibration. The C_2H_4 concentration prediction errors are calculated with and without shift correction. The error has been

reduced from 6.92% to 1.50% by ELM method after shift correction. There is no noticeable error improvement with peak tracking and CLS method. The errors remain in 20% range.

In a summary, the optimized ELM model has demonstrated dramatic improvement of TDLAS prediction accuracy by one order of magnitude in the trace C_2H_4 measurement in a simulated industrial process analysis scenario. The robustness enhancement is explicit. ELM algorithm could be extended to other areas of spectra analysis in the harsh application environments and dynamic process background. Complicated spectra deformation in addition to spectra shift could be dealt with ELM algorithm for future research.

Author Contributions: Conceptualization, Wenhai Ji and Guolin Li; methodology, Wenhai Ji and Li Zhong; software, Ying Ma and Xiaocui Lv; validation, Di Song; formal analysis, Ying Ma; investigation, Li Zhong; writing—original draft preparation, Wenhai Ji.; writing—review and editing, Chuantao Zheng and Guolin Li.

Funding: This research received funding from Natural Science Foundation of China (Nos. 61775079, 61627823), Natural Science Foundation of Shandong Province (No. ZR2017LF023), Key Science and Technology R&D program of Jilin Province (No. 20180201046GX), Huimin Special Project of Qingdao Science and Technology Bureau (No. 17-3-3-89-nsh), Jilin University State Key Laboratory on Integrated Optoelectronics Open Research Fund Project (No. IOSKL2017KF0).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Y. Wei.; J. Chang.; J. Lian.; *et al.* A Coal Mine Multi-Point Fiber Ethylene Gas Concentration Sensor. Photonic Sensors 2015, 5(1), 67-71.
2. D. Shi.; W. Song.; J. Ye.; *et al.* Experimental Investigation of Reacting Flow Characteristics in a Dual-Mode Scramjet Combustor. International Journal of Turbo & Jet-Engines 2016, 33(2), 95-104.
3. Y. Wang.; Y. Wei.; J. Chang. Tunable Diode Laser Absorption Spectroscopy Based Detection of Propane for Explosion Early Warning by Using a Vertical Cavity Surface Enhanced Laser Source and Principle Component Analysis Approach. IEEE Sensors Journal 2017, 17, 4975-4982.
4. J. Jiang.; M. Zhao.; G.-M. Ma.; H.-T. Song.; C.-R. Li.; X. Han.; C. Zhang. TDLAS-Based Detection of Dissolved Methane in Power Transformer Oil and Field Application. IEEE Sensors Journal 2018, 18, 2318-2325.
5. Z. Yuan.; X. Yang.; W. Xie.; X. Li. Research on the Online Test of Diesel NOx Emission by TDLAS. Spectroscopy and Spectral Analysis 2018, 38, 194-199.
6. H. Li.; Y. Zhu.; F. Dong.; *et al.* Simulation Design and Verification of CO Monitoring Based on Tunable Diode Laser Absorption Spectroscopy. Precision Electromagnetic Measurements (CPEM) 2010 Conference on. IEEE, 2010, 482-483.
7. Y. Wang.; Y. Wei.; T. Liu.; *et al.* TDLAS Detection of Propane/Butane Gas Mixture by Using Reference Gas Absorption Cells and Partial Least Square Approach. IEEE Sensors Journal 2018, 18(20), 8587-8596.
8. P. Barmpalexis.; A. Karagianni.; I. Nikolakakis.; K. Kachrimanis. Artificial Neural Networks (ANNs) and Partial Least Squares (PLS) Regression in The Quantitative Analysis of Cocrystal Formulations by Raman and ATR-FTIR Spectroscopy. Journal of Pharmaceutical and Biomedical Analysis 2018, 158, 214-224.
9. C. Malegori.; E. J. Nascimento Marques.; S. T. de Freitas.; M. F. Pimentel.; C. Pasquini.; E. Casiraghi. Comparing the Analytical Performances of Micro-NIR and FT-NIR Spectrometers in the Evaluation of Acerola Fruit Quality, Using PLS and SVM Regression Algorithms. Talanta 2017, 165, 112-116.
10. B. Le.; D. Xiao.; Y. Mao. Coal. Classification Based on Visible, Near-Infrared Spectroscopy and CNN-ELM Algorithm. Spectroscopy and Spectral Analysis 2018, 38, 2107-2112.
11. Y. Yang.; S. Zhang.; Y. He. Dynamic Detection of Fresh Jujube Based on ELM and Visible/Near Infrared Spectra. Spectroscopy and Spectral Analysis 2015, 35, 1870-1874.
12. T. Owolabi.; M. Gondal. Development of Hybrid Extreme Learning Machine Based Chemometrics for Precise Quantitative Analysis of LIBS Spectra Using Internal Reference Pre-Processing Method. ANALYTICA CHIMICA ACTA 2018, 1030, 33-41.

13. Y. Chen.; Z. Wang.; Z. Wang.; X. Li. Research on Concentration Retrieval of Gas FTIR Spectra by Interval Extreme Learning Machine and Genetic Algorithm. Spectroscopy and Spectral Analysis 2014, 34, 1244-1248 .
14. Y. Bu.; G. Zhao.; J. Pan.; Y. Bharat Kumar. ELM: an Algorithm to Estimate the Alpha Abundance from Low-resolution Spectra. Astrophysical Journal 2016, 817, 78.
15. X. Bian.; C. Zhang.; X. Y. Tan.; M. Dymek.; Y. Guo.; L. Lin.; B. Cheng.; X. Hu. Boosting Extreme Learning Machine for Near-Infrared Spectral Quantitative Analysis of Diesel Fuel and Edible Blend Oil Samples. Analytical Methods 2017, 9, 2983-2989.
16. G. Huang.; Q. Zhu.; C. Siew. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. In Proceedings of the IEEE International Conference on Neural Networks, Budapest, Hungary, 25-29 July 2004; 985-990.
17. G. B. Huang.; Q. Y. Zhu.; C. K. Siew. Extreme Learning Machine: Theory and Applications. Neurocomputing 2006, 70, 489-501.
18. G. B. Huang.; D. H. Wang.; and Y. Lan. Extreme Learning Machines: A Survey. International Journal of Machine Learning & Cybernetics 2011, 2, 107-122.
19. G. B. Huang. An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels. Cognitive Computation 2014, 6, 376-390.
20. G. B. Huang.; H. Zhou.; X. Ding.; R. Zhang. Extreme Learning Machine for Regression and Multiclass Classification. IEEE Trans Syst Man Cybern B Cybern 2012, 42, 513-529.
21. X. X. Yin.; S. Hadjiloucas.; J. He.; Y. Zhang.; Y. Wang.; D. Zhang. Application of Complex Extreme Learning Machine to Multiclass Classification Problems with High Dimensionality. Digital Signal Processing 2015, 40, 40-52.
22. Z. Huang.; Y. Yu.; J. Gu.; H. Liu. An Efficient Method for Traffic Sign Recognition Based on Extreme Learning Machine. IEEE Transactions on Cybernetics 2016, 47, 920-933.
23. R. J. Kowalski.; C. Li.; G. M. Ganjyal. Optimizing Twin-Screw Food Extrusion Processing through Regression Modeling and Genetic Algorithms. Journal of Food Engineering 2018, 234, 50-56.
24. W. D. Pan.; J. W. Zhang.; J. M. Dai.; Y. F. Zhang. Tunable Diode Laser Absorption Spectroscopy for Simultaneous Measurement of Ethylene and Methane near 1.626 μm. Journal of Infrared & Millimeter Waves 2013, 32, 486-490.
25. F. V. D. Schoor.; F. Verplaetsen. The Upper Explosion Limit of Lower Alkanes and Alkenes an Air at Elevated Pressures and Temperatures. Journal of Hazardous Materials 2016, 128, 1-9.
26. S. W. Sharpe.; T. J. Johnson.; R. L. Sams.; P. M. Chu.; G. C. Rhoderick.; P. A. Johnson. Gas-phase Databases for Quantitative Infrared Spectroscopy. Applied Spectroscopy 2004, 58, 1452-1461.
27. K. Javed.; R. Gouriveau.; N. Zerhouni. SW-ELM: A Summation Wavelet Extreme Learning Machine Algorithm with A Priori Parameter Initialization. Neurocomputing 2014, 123, 299-307.
28. F. Han.; M. R. Zhao.; J. M. Zhang.; Q. H. Ling. An Improved Incremental Constructive Single-Hidden-Layer Feedforward Networks for Extreme Learning Machine Based on Particle Swarm Optimization. Neurocomputing 2016, 228,133-142.
29. C. Yeom.; K. Kwak. Short-Term Electricity-Load Forecasting Using a TSK-Based Extreme Learning Machine with Knowledge Representation. Energies 2017, 10, 1613.