Volume 8 Issue 2, 2022
DOI: 10.6919/ICJE.202202_8(2).0008

Carbon Monoxide Sensor Correction based on Local Linear Regression

Xingyu Cai[1,*], Xiaowen Sun[2,a], Ziwen Sun[2,b]

[1] Jiangnan University, Wuxi, Jiangsu, China

[2] Colorado State University, CO, US

[a]sun.sxw@foxmail.com, [b]danicas@colostate.edu

[*]Corresponding author: 1024101233@qq.com

These authors contributed equally to this work

Abstract

Drift is one of the most serious impairments afflicting gas sensors. It can be seen as a gradual change in the sensor response over a long period of time when the external conditions are constant. This paper proposes a simple correction method based on local linear regression. Use the original data as a fitting plane to correct the drift of the sensor. Using linear regression and local linear regression to process the original data respectively, it is found that both methods are effective and the local linear regression effect is better.

Keywords

Gas Sensors; Drift Correction; Local Linear Regression.

1. Introduction

1.1 Atmospheric Environment Situation

The atmosphere is made up of many trace gases whose dynamics directly or indirectly influence environmental and climate change, mainly dominating the chemical balance in the atmosphere. The concentrations of these gases are significantly affected by human activities.

Carbon monoxide (CO) is a colorless and odorless gas. CO easily combines with hemoglobin in the blood to form carboxyhemoglobin, which causes the hemoglobin to lose its ability and function to carry oxygen, resulting in tissue asphyxiation and, in severe cases, death.

The main sources of carbon monoxide (CO) in the atmosphere are the combustion of fossil fuels and biomass and the majority of CO affecting urban air is from exhaust emissions from motor vehicles and incomplete combustion of fossil fuels and natural gas in daily life. Long-term exposure to low concentrations of CO can be hazardous to human health and contribute to global warming, therefore, it is important to pay attention and use sensitive CO gas sensors for monitoring.

1.2 Current Technology Status of Atmospheric Environment Testing

Atmospheric environmental monitoring is a way to measure the types and concentrations of pollutants in the atmosphere. The purpose is to identify pollutants in the atmosphere through monitoring, summarize their distribution characteristics and dispersion patterns, so as to strictly control the emissions of air pollution sources, and provide an accurate basis for environmental management decisions. It reflects the impact and hazard of various air pollutants on-air environment quality. The monitoring of air pollution sources in China started late compared with foreign countries, and the means and equipment used have a certain lag. As the number of pollutants with complex composition

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0008

and difficult to handle in the atmosphere increases, the traditional monitoring methods and means can no longer accurately reflect the environmental quality and change trends, which makes it difficult for researchers to understand, analyze and study environmental pollution problems. Therefore, to further improve or optimize the current means of monitoring the air environment in China and to improve the level of monitoring technology is a subject that needs to be studied urgently.

1.3 Application of Sensors and the Purpose of the Research

At present, more and more attention is paid to the monitoring of gases all over the world, and higher requirements are put forward for the monitoring of toxic and hazardous gases, including NO_x, SO_2, CO_x in the atmospheric environment. Commonly available gas sensors are electrochemical sensors, metal oxide semiconductor type sensors, etc. Metal oxide semiconductor sensors mainly use the adsorption effect on the air to be measured, and the corresponding display is generated by the mutual contrast of the change of current. Nevertheless, such kinds of sensors have many drawbacks, and when the surrounding environment changes, the measurement results of the sensor will produce a large error. Therefore, the research and development of low-power, low-cost, miniaturized, high-sensitivity, and high-stability gas sensors, especially carbon monoxide sensors, is of great importance to protect human health and protect the ecological environment. With the progress of research, atmospheric monitoring using carbon monoxide sensors has become one of the current important topics.

There are two main measurement methods for the sensor in this experiment, the long-range method, and the sensor method. The long light range method has the characteristics of high detection accuracy, high cost of the instrument, and long detection period, and normally requires professional personnel to operate. In the sensor method, the metal oxide semiconductor and electrochemical sensors mainly used in the experiment have lower costs and higher accuracy, however, they are easily affected by temperature and humidity, drift phenomenon, and unstable consistency of the same batch of sensors. Because the calibration gas is more expensive and difficult to transport, using the data from the universal monitoring station as the real data and the sensor data for the data to be corrected can be better and more convenient for data processing and comparison.

2. Datasource

The study of air pollution in Beijing is of great significance to the development of the environment, however, there are few studies on the distribution characteristics of air pollution in Beijing and the influence of economic development on air pollution. The research using CO sensors can help to understand the current situation of air pollution management in Beijing and propose the following management policies from the perspective of economic development. Therefore, this study is more relevant, targeted and can complement the existing studies with certain theoretical significance. In addition, this paper can provide theoretical and data support for the government's decision making, and put forward specific suggestions on air pollution control in Beijing, which has practical significance. In order to avoid the continued deterioration of CO impact on people's lives in the city, we will take the atmospheric monitoring of Haidian District in Beijing in 2017 as an observation, and propose solutions to improve air quality by using sensors of CO to monitor the CO content in the air, selectively interact with the substances to be measured, and transform the measured chemical parameters into information that the conduction system can produce a response.

In order to monitor air quality, the state has set up air quality monitoring stations to sample, measure, and analyze pollutants in the atmosphere and air at fixed points, continuously or at regular intervals. In order to monitor the air, several air stations are generally set up in a critical city of environmental protection. Multi-parameter automatic monitoring instruments are installed in the stations for continuous automatic monitoring, and the monitoring results are stored in real-time and analyzed to get relevant data. Generally speaking, the monitoring methods are the long optical path method and sensor method. Although the long optical path method is highly accurate, its cost is high, and the temperature required for the environment is also high. Generally, it only exists in large cities, so it is

Volume 8 Issue 2, 2022
DOI: 10.6919/ICJE.202202_8(2).0008

not discussed here. The cost of the sensor method is low, and it exists in many urban air monitoring stations, but the accuracy of air monitoring is insufficient. In order to improve the accuracy of the sensor, it has more popularization significance. The CO monitoring data of Wanliu Monitoring Station in Haidian District, Beijing, are used to correct the sensor data. Reference to the accurate data collection method of Wanliu monitoring station, that is, using urban encrypted grid point measurement or model simulation calculation method, estimate the overall average value of pollutant concentration in the city's built-up area, which is of more reference significance for sensor correction.

3. Technology

The data processing steps are as follows.

(1) Extracting CO data from Wanliu, Beijing.

Table 1. Assuming that date is the time of the observed data, hour represents the time of presence, and wl_CO denotes the carbon monoxide concentration in Wanliu, Beijing.

date	hour	wl_CO
201710001	1	1.7

(2) Extract the CO data collected by the actual sensor.

Table 2. Assuming that date is the time of the observed data, hour represents the time of presence, sr_CO denotes the sum of carbon monoxide divided by number, T and H indicate temperature and humidity, respectively.

date	hour	sr_CO	T	H
20171001	1	1.5	0.3	0.2

1) The time is limited between 00 and 01 points.

2) All the CO, T, and H data in this time period are read out and summed up by the corresponding numbers, CO_SUM, CO_NUM.

3) sr_CO=CO_SUM/CO_NUM.

(3) Integrate the data from Wanliu, Beijing, and the actual sensor accordingly.

Table 3. Assuming that date is the time of the observed data, hour represents the time of presence, sr_CO denotes the sum of carbon monoxide divided by number, T and H indicate temperature and humidity, respectively, and wl_CO denotes the carbon monoxide concentration in Wanliu, Beijing.

date	hour	Sr_CO	T	H	wl_CO
20171001	1	1.5	0.3	0.2	1.7

(4) Turn the data set into the form of data.txt and input it to the locally weighted regression algorithm.

(5) Calculate the error.

The iterative solution analysis algorithm of the K-means clustering algorithm mainly divides the data into K groups, randomly selects K objects as the initial cluster centers, and then calculates the distance between each object and each cluster center, assigning each object to the cluster center nearest to it. With each assignment, the cluster centers are recalculated based on the existing objects in the clusters,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202202_8(2).0008

and this process is repeated until a termination condition is met. Parameter optimization (K) is based on the error between the fitted and true values.

We calculate the root mean square error of the test samples according to different K values separately and select the optimal K value according to the error performance.

The specific steps of the experiment are as follows.

First and foremost, we collect and process the data from Wanliu, Beijing, obtain and analyze the actual sensor data through experiments, and present them in separate tables. Then, we combined the data, divided them into training and test data (4: 1), calling the algorithm to calculate and analyze the error tuning parameters.

4. Results

In order to verify the effectiveness of the local linear regression model in practical applications, conduct carbon monoxide detection experiments. The data from the Wanliu atmospheric monitoring point in Haidian District, Beijing was used to correct the carbon monoxide sensor installed near Wanliu. The set sensor collects the content, temperature and humidity of various components in the atmosphere every one second. Take the carbon monoxide content under different temperature and humidity collected 24 hours from July to October as samples. Use the atmospheric carbon monoxide content released every hour from July to October by the Wanliu Testing Station to calculate the deviation value and calibrate the sensor.

4.1 Data Processing

Table 4. Normalized real data of Wanliu monitoring station, wl_co is the normalized result of carbon monoxide concentration collected from Wanliu

date	hour	wl_co
201779	1	0.4
201779	2	0.32
201779	3	0.28
201779	4	0.28
201779	5	0.32
201779	6	0.36
201779	7	0.36
201779	8	0.36

Table 5. Normalized sensor sample set

date	hour	sr_co	temperature	humidity
201779	1	0.4911	0.5089	0.8705
201779	2	0.5182	0.5121	0.8791
201779	3	0.5064	0.5124	0.8741
201779	4	0.4933	0.5144	0.8705
201779	5	0.4876	0.5147	0.8744
201779	6	0.4915	0.5171	0.8865
201779	7	0.5051	0.5226	0.8800
201779	8	0.5208	0.5324	0.8712

 Volume 8 Issue 2, 2022
DOI: 10.6919/ICJE.202202_8(2).0008

The hourly carbon monoxide content was intercepted from the data released by the Wanliu Monitoring Station, and normalize the temperature and humidity of carbon monoxide concentration. Some of the data sets formed are shown in Table 4. The carbon monoxide content, temperature, and humidity measured by the sensor are averaged hourly, and normalize the temperature and humidity of carbon monoxide concentration. Some of the formed data set is shown in Table 5. Combine the two tables into one and input them into the local linear regression and linear regression algorithms respectively.

4.2 Linear Regression

Take temperature as the second column of the matrix, humidity as the third column, and fill 1 in the first column as the matrix X, and the carbon monoxide content detected by the sensor is taken as the result matrix Y. First, the matrix X and the transpose of the matrix X are multiplied to obtain the matrix xTx. If xTx is an invertible:

$$xTx=XTX \tag{1}$$

matrix, it is multiplied by the transpose of matrix X and matrix Y, and the resulting matrix is the coefficient matrix θ. Use the coefficients of the matrix θ into the resulting equation:

$$\theta=xTx-1XTY \tag{2}$$

to get the corrected Y’ matrix.

$$Y'= \theta X \tag{3}$$

Local linear regression.

The method of obtaining the values of matrix X and matrix Y is the same as that of linear regression. But the matrix θ is calculated differently.

$$\theta =(XTWX)-1XTWY \tag{4}$$

W is a weight matrix containing only diagonal elements. Xi is the sample point, X is the:

$$W(i,i)=\exp(\frac{(X_i-X)^2}{-2k^2}) \tag{5}$$

prediction point.The result matrix Y’ is obtained in the same way as linear regression.

Error calculation.

Table 6. Normalized real data of Wanliu monitoring station

method		square difference
Uncorrected		73.9559
After correction	linear regression	45.2716
	local linear regression	43.7330

ISSN: 2414-1895
DOI: 10.6919/ICJE.202202_8(2).0008

The square difference of the data is used to calculate the error. The square difference of sensor data before correction, the square difference after correction by linear regression algorithm and the square difference after correction by local linear regression algorithm are shown in Table 6.

4.3 Result Analysis

Take 1611 points and use local linear regression to calculate the fitted a curve on a figure. In order to enhance the persuasiveness, the curve calculated by the linear regression and normalized true carbon monoxide concentration is also drawn on the same figure. The comparison between them is shown in Figure 1.

Through comparison, it can be seen that the accuracy of the carbon monoxide sensor is greatly improved after correction using the local linear regression algorithm. At the same time, it can be seen that the local linear regression algorithm is better than the linear regression algorithm, which can greatly improve the accuracy of the sensor.

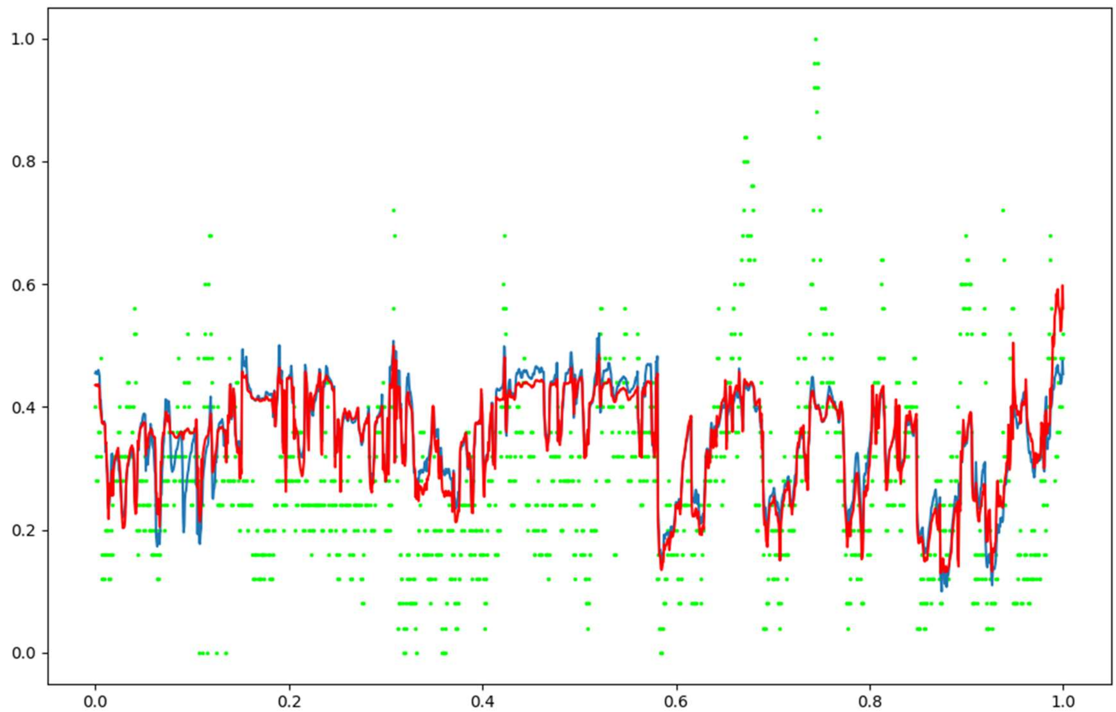

Figure 1. Comparison among linear regression, local linear regression and normalized true carbon monoxide concentration, red curve is the result of linear regression,blue line is the result of local linear regression, green points are true data.

5. Conclusion

In this paper, in order to verify the effectiveness of the local linear regression model in practical applications, we conducted a carbon monoxide detection experiment using data from an atmospheric monitoring site in Gulf Stream, Haidian District, Beijing, to calibrate a carbon monoxide sensor installed near Gulf Stream. The experimental results show that our calibration using the local linear regression algorithm dramatically improves the accuracy of the carbon monoxide sensor. It can also see that the local linear regression algorithm is superior to the linear regression algorithm, which can significantly improve the sensor's accuracy. To further enhance the calibration performance, we can detect and remove anomalies before the sensor calibration process, thus allowing us to simulate potential physical phenomena and detect anomalies. The calibration method is based on stepwise regression. In the future, we can continue to learn this method, which automatically and systematically selects the appropriate support parameters for the calibration function. The evaluation of anomaly detection shows that the results outperform state-of-the-art methods in terms of accuracy, precision, and completeness. The removal of anomalies before calibration can further improve data quality.

References

[1] Yao S.H. He D.Q. Zhou C.P. Lu Y.Q. Hao J.M. Atmospheric variation of NOx, CO, and O3 in Beijing [D]. Study on the variation pattern of NOx, CO and O3 in Beijing city atmosphere [D]. Beijing: Department of Environmental Science and Engineering, Tsinghua University, 1999 1-4.

[2] Miao Lei, Beijing PM2.5 concentration evolution characteristics and meteorological conditions influence analysis [D]. Beijing: China Academy of Meteorological Sciences,2016 1-83.

[3] Wu Hongyi1 Wang Pucai1 Wang Gengchen1 Zheng Youfei2 Kong Qinxin Variation of total atmospheric CO column in Beijing in autumn[A]. Beijing: Department of Middle Atmosphere and Remote Sensing Theory and Technology Research, Institute of Atmospheric Physics, Chinese Academy of Sciences, Nanjing:Nanjing Meteorological Institute, 2005 1-8.

[4] Sun, J.L. Liu, D.M. Yang, X. Atmospheric particulate matter pollution levels and its influencing factors in Haidian District, Beijing[A]. Beijing: School of Energy, China University of Geosciences, 2009 1-5.

[5] Yin Xiaodi. Study on the impact of economic development on air pollution in Beijing-Tianjin-Hebei region [D]. Hebei: Hebei University, 2017 1-58.

[6] Li Xiangqin. Research on the Application of Big Data Analysis Technology in Atmospheric Environment Monitoring[J]. China New Technology and New Products, 2021(3): 122-124.

[7] Wu Tianyu. Research on the application of ground-air integrated atmospheric environment monitoring technology[J]. Industrial & Science Tribune, 2017, (16) 23: 42-43.

[8] Hou Yuexi. Application of Carbon Monoxide[J]. MECHANICAL MANAGEMENT AND DEVELOPMENT, 2017, (10): 87-88, DOI: 10.16525/j.cnki.cn14- 1134/ th.2017.10.40.

[9] Yang Bangchao, Duan Jianhua. Application and advance of CO sensors[J]. 2001, Vol. 20 (12):1-4.

[10] Kundu, S., Sudharson, R., & Narjinary, M. PD impregnated gallia: Tin oxide nanocomposite-An excellent high temperature carbon monoxide sensor[J]. Sensors and Actuators B: Chemical, 2018(254), 437-447.

[11] Qiao, S., Ma, Y., He, Y., Yu, X., Zhang, Z., & Tittel, F. K. A sensitive carbon monoxide sensor based on Photoacoustic spectroscopy with a 2.3 μm mid-infrared high-power laser and enhanced gas absorption[J]. Harbin, China, Rice University, USA: Sensors, 2019, 19(14), 3202.